In past articles, Jason Weiss and I have discussed database design and security as they relate to I-O psychology (Weiss & Worst, 2002; Worst & Weiss, 2003). Here, I will discuss how databases are being used to leverage organizational data in a family of software applications known collectively as Business Intelligence (BI).

BI is a fairly new term that incorporates a broad variety of processes and technologies for harvesting and analyzing information to help businesses make decisions. In the service of this goal, business intelligence software may include such activities as six-sigma, statistical analysis, data mining, and financial forecasting. A wide array of companies, both new and old (e.g., Microstrategy and Microsoft), are now developing tools dedicated to providing unified BI functionality. In fact, some companies with BI products are very well known in the I-O community (Read on to find out who they are!).

I believe it is important for I-O psychologists to have a general understanding of this area because it may impact the practice of I-O in the near future. How? As I’ve discussed in the past, most organizations are well on the way to moving their HR data from file drawers to electronic databases stored on central servers. As a result, data that were formerly difficult to access and leverage have become available assets, waiting to be leveraged with BI tools. BI will enable us to answer questions that we were previously unable to approach. As a cautionary note, it could potentially enable others to use sensitive HR data to answer questions that they might not have the background to consider. We, therefore, must consider our role with respect to BI as dualy focused on leveraging its strengths for the advancement of our field but protecting others from potentially dangerous misuse of the technology.

BI is still a work in progress, so there is no widespread agreement as to the minimal functionality of a BI tool. In this article, I will describe four types of functionality commonly found in BI software packages. At the end, I’ll consider examples of how they might come into play in I-O.
Enterprise Reporting and Report Distribution

Developing relevant, valid, and timely business information is a key purpose of BI. However, that’s only half the battle; the assembly of critical information is useless if it isn’t delivered to the right people in a timely manner. Consider for a moment what might be called the status quo of information creation and distribution: For over 40 years, organizations have run standardized reports that are distributed to key information consumers. These reports were run by the IT department as batch jobs at night. After the reports were generated, they then had to be printed and distributed—often by mail—taking a significant amount of paper and time to print and distribute. Using this process, a report might not find its target until several days or a week after it was initially generated.

Report generation and distribution today is very different. Now reports can be run at any time of the day or night and immediately distributed via e-mail or simply put on a corporate intranet, greatly reducing the cost of report generation and distribution. More flexible and user-friendly report generation tools mean that a greater variety of reports can be made available to users. In addition, as reporting needs have gone from domestic to international, some software has made an additional leap in flexibility to generate reports in multiple languages and metrics. For example, some packages can generate a sales report for a multinational company in English using the U.S. dollar and in Japanese using yen.

The primary purpose of enterprise reporting is to automatically generate and distribute standardized reports to management at specified intervals. For example, recruiting information broken down by region may be delivered to the VP of HR each week while regional HR managers may get local reports daily. This enables organizations to respond quickly to both opportunities and competitive pressures by giving managers and other knowledge workers fast, easy access to key information to make rapid front-line decisions.

Here is a sample of the report delivery functionality contained in many BI products:

• One-click distribution of a large number of reports to people inside and outside the enterprise;
• Automatic report distribution on a predetermined schedule;
• On-event distribution based on predetermined events triggering a report distribution (e.g., unusual drop in inventory); and
• Self-subscription or administrator-based subscription to a report(s).

These are just examples of the types of functionality in BI software to make sure that people who need information have access right when they need it.

Cube Analysis

In our past discussions of databases, I have described a data table as hav-
ing two axes: the horizontal axis (rows) defined each record and the vertical axis (columns) or fields defined the data in each record. For example, a data table containing the first name, last name, and telephone number for 100 people would have 100 records and three fields. These types of two-dimensional data views are probably the ones most commonly used by I-O psychologists when analyzing experimental data on subjects.

Over the past decade or so, a new approach for organizing data has emerged called On Line Analytical Processing (OLAP). Although OLAP data can be viewed in a two-dimensional format, it is more often viewed with three or more axes in the form of a cube. In fact, there is theoretically no limit to the number of axes an OLAP cube can have, creating a hypercube. Below are some examples of OLAP cubes. Note that each cube is a subset of highly interrelated data that is preorganized to allow users to combine any attributes in the cube (e.g., stores, products, customers, suppliers) with any metrics in the cube (e.g., sales, profit, units, age) to create various views that can be displayed on a computer screen.

Cube analysis is used most often by people like managers who have a deep interest in understanding the root causes underlying the data in reports but do not possess skills for full ad hoc investigation of the databases. Cube analysis lets people flip through a series of report views, using the now standard OLAP features of:

- Rotation to new dimensional comparisons in the viewing area (i.e., rotate from a customer/region view to a customer/product view).

Analysis Cubes

- Investment in Asset as % Revenue
- Operating Cash by BU
- Working Capital LO QTQ
- Actual vs. Plan
- Actual vs. Forecast
- Actual vs. Budget
- Holding Trends by Season
- Performance by Product
- Performance by Region
- Sales by Customer
- Return Trends by Store

• Slicing subsets for on-screen viewing (e.g., view only a slice of a cube such as all products and customers in one region).
• Drill down to deeper levels of consolidation (i.e., make the cube smaller to expedite searches and data analyses).
• Reach through to underlying detail data. Reach through is a means of extending the data accessible to the end user beyond that which is stored in the OLAP server. A reach through is performed when the OLAP server recognizes that it needs additional data and automatically queries and retrieves the data from a data warehouse or OLTP system.

All of these OLAP features, first introduced in the early 1990s, allow users to slice and dice a cube of data, or analysis cube, using simple mouse clicks.

**Ad-Hoc Query and Analysis**

Although standardized reports are useful for many users, they can not meet the informational needs of users who need more flexible data access. Practically speaking, it would be impossible to predesign all possible report permutations because this would involve thousands—if not millions—of combinations. Under the old, “centralized” process, all requests for “special” reports and analysis had to go through the corporate IT department, if they were permitted at all.

With the advent of client/server computing, new interactive software tools enable managers to generate their own ad-hoc analysis and reports. For example, standardized reports may require a manager to wade through vast amounts of data to find some needed piece of information. Ad-hoc analysis and reporting tools enable managers to identify not only the exact information they need but also to perform specific analyses of the data they need.

Although at first glance this seems to make imminent sense, experience has shown that managers sometimes still have trouble using the tools for creating and formatting reports that come with the BI software. These tools are really not intended for use by everyone in the organization; rather, they should be leveraged by those with some data analysis background and higher levels of software skills. These users will need to conceptually understand how to filter data into subsets, perform analyses on those subsets, and then generate formatted reports containing various types of charts or graphs. More sophisticated software will use a variety of prompts, wizards, and other aids to help accomplish these tasks, but the user still has to understand the data and which analysis approach is appropriate for the conclusions being drawn.

**Statistical Analysis and Data Mining**

Where ad-hoc analysis is intended for those with average analytical skills, this area of BI is the most sophisticated and targeted at professional information analysts who daily perform correlation analysis, trend analysis, and pro-
jections. In addition, extremely sophisticated modeling tools such as neural networks may also be part of the data mining package. These analysts apply mathematical, financial, and statistical functions against the data of the entire enterprise. Although the tools now require little to no programming skills, they still need a very skilled user who will know which statistical and data mining techniques can provide the proper and best insight into a problem or business question.

Most people reading this article will know that both SAS and SPSS (companies I came to know and love during my undergraduate/graduate days) develop very sophisticated data analysis tools. But did you know that both of these companies have soup-to-nuts, integrated BI tools? SPSS’s product is called ShowCase Suite and includes all of the features discussed in this article. SAS’s product is called SAS Enterprise BI Server and also includes all of the features discussed including OLAP, ad-hoc reporting and analysis, and so forth. Both products have a single integrated interface that can be used for implementing all of their BI tools for data analysis, report development, and report distribution. In addition, a single unified database provides the foundation upon which all of these tools rest.

A Sample Application of BI in I-O

Suppose the head of HR for a large network of hospitals has just received a report generated by the report distribution engine showing that the hospital network is well behind its hiring goals for nurses. After running some canned OLAP cube analysis reports, she learns that it’s not all nurses but primarily RNs that are far behind hiring goals. She then runs some ad-hoc analyses looking at how many people applied for RN positions over the past quarter, how many were offered jobs, and how many accepted. She compares the past quarter with the same quarter a year ago. She learns that the proportion of applicants to those given job offers is basically the same as last quarter but acceptance of offers is significantly lower. She then wonders if this is affecting the quality of nurses being hired, so she uses the data mining tools to compare validated employment test scores from a year ago versus those from the past quarter. She notices a significant decrease even though both groups meet the minimum cut scores. After pondering why this might be happening, she conducts some more ad-hoc analyses to see if there is an increase in the time between job application and job offer. She notices that this time has increased by 4 weeks, on average, and hypothesizes that possibly the best candidates have already taken offers from other hospitals by the time they get the job offer from her hospital network. She meets with her senior staff, presenting the above findings, and tasks them with calling a sample of candidates who rejected job offers to find out why and to decrease the amount of time it takes to go from application to job offer/rejection.
BI and I-O: Next Steps

Human resource data is an important corporate asset that will be utilized much more in the future for making business decisions. This will place a difficult burden on most HR departments because their employees are not trained or experienced in the use of analytical tools and techniques. I-O psychologists have some of the best training and experience for conducting HR data analysis, and we are very familiar with the data and the issues surrounding that data (e.g., employment application and test results). Although many of the tools in BI software packages are familiar to most of us, additional training would be needed for some tools such as OLAP and neural networks, but that shouldn’t be too difficult for datacentric I-O types to master.

The availability of much better organized corporate HR data and powerful analytical tools will enable us to investigate questions that we might not have ever considered in the past. We will need to expand how we think about HR data analysis because data that was highly fragmented and difficult to access in the past will soon be consolidated into one large, organized, and easily accessible corporate data warehouse. For example, HR will be able to conduct OLAP reach-through analyses from the HR data cubes over to performance data contained in OLAP cubes for the sales department. In addition, this new capability will enable us to reopen old questions that we may feel need further investigation.

Those of us working in or consulting to large organizations may need to do some investigation as to whether or not there is an effort underway to building a consolidated BI database. The BI effort may well be underway but has not been communicated to everyone who needs to know about it. If you find that one of these efforts is underway, or was completed a while ago, let the appropriate people know that you would like to receive training on how to use the software package and tools. It should be easy to justify how you could use the data to improve your own job performance and that of your department. In addition, as we’ve discussed in previous articles, you should be aware of what security measures are being taken to protect the HR data and decide if you feel they are adequate. See Worst & Weiss (2003) for more information on this very important topic.

Time To Be Proactive

The development and use of BI tools is well on its way, and it is inevitable that someone who is not trained to work with HR data will incorporate it into BI analyses. We should not wait for time to tell if the people doing these analyses and making recommendations as a result are doing so responsibly. We can influence the “outside” use of HR data by being aware of the growing availability of these data for analysis and then by asserting our roles as custodians of the data. Otherwise, we’ll only have ourselves to blame as we wince at decisions based on misunderstood data and/or faulty analysis.
Web Resources

There is a vast array of information about BI on the Internet. I did some of the leg work for you and identified several good BI Web sites below. I can be reached at kensei@comcast.net if you have any questions.

http://www.dmreview.com/corporate/aboutus.cfm
DM Review is a business intelligence, analytics, and data warehousing publication. A great place to start.

http://www.businessintelligence.com/index.asp
BusinessIntelligence.com has a broad variety of BI articles including BI research. It also has forums where you can discuss various BI topics.

http://www.computerworld.com/databasetopics/businessintelligence/story/0,10801,93940,00.html
Interesting Computer World article (6/04) on the future of BI based on the opinions of industry leaders.

http://businessintelligence.ittoolbox.com/
A broad variety of articles on BI along with some information on ITToolbox’s own products. There’s even some information on careers in BI.

www.microstrategy.com/
Microstrategy, large vendor of BI software.

References


Bibliography

Microstrategy (2002). The 5 Styles of Business Intelligence.