+ + tableau®
SOFTWARE

fast analytics and business intelligence for everyone:
best practices for deploying collaborative bi

AUTHOR:

Neil Raden
Co-founder of
Smart (enough) Systems, LLC

DATE:

May 4, 2008

EXECUTIVE SUMMARY

Software for Business Intelligence (BI) provides the mechanism for people in organizations to examine historical trends, to measure results, to look for patterns in the data that might be missed in standardized reports and event to create alerts that fire when certain conditions are or aren't met. But the BI industry has one intractable problem it has yet to solve — how to actually get people to use it.

BI tools for the most part lack collaborative capabilities and they especially lack visualization. When you consider that even a simple dimensional model of a business area may contain a hundred attributes in ten or twelve dimensions with hierarchies, it is easy to see how an exhaustive catalogue of reports of every combination and rotation could easily exceed one million variations. This limitation of two-dimensional row and column reports, is lifted with the addition of not only visualization, but interactive visualization.

Yet, BI adoption rates are low and BI practice methodologies are out of date. They fail to take into account that useful data for analytics is found in many places beside the data warehouse, that younger people entering the workforce are inclined to use tools that are internet-related, not enterprise and that people don't live in pyramids . A power user may be a novice in some areas and vice versa. Roles are defined too tightly and capabilities are unfortunately deployed against this faulty model.

Most important of all, being informed is useful, but it is making the right decisions that counts. That takes more than reports and dashboards.

Classic BI is a combination of "push and pull" (some information is pushed out to users by the system, some information needs to be pulled out by the user), but its scope is usually the past up to almost the present. Decisions that once might have taken days now have to be made at the speed of the transaction, such as while your customer is completing an online transaction. That is impossible if it requires submitting a query, or even looking at a dashboard that is updated every few minutes. Analytical capabilities have to inform the decision-making process the same way embedded software will inform a mission to Mars — there just isn't time to analyze, ponder and react. Instead, people need answers to questions instantaneously; analytical tools need to provide input and evaluative capabilities. For certain kinds of decisions, models have to make the decisions and their decisions will be only as good as the analytics that went into their formation.

Today's organizations are complex webs of interacting groups and projects often created on-the-fly with demands for near-immediate results. Providing a systematic means for users to draw rapid, actionable insights and then share those insights is a constant challenge. This paper will show you how to navigate around the usual hurdles and deliver collaborative BI that works.

MOVING FROM REPORTS TO DECISIONS

The late Herbert Simon, a Nobel Prize laureate in Economics, wrote extensively about decision-making and the role of analytics in organizations. Simon considered problem solving and decision making separate parts of the same process:

> *"…making decisions and solving problems. It is work of choosing issues that require attention, setting goals, finding or designing suitable courses of action, and evaluating and choosing among alternative actions. The first three of these activities — fixing agendas, setting goals, and designing actions—are usually called problem solving; the last, evaluating and choosing, is usually called decision making."*

All of these activities involve working both individually and in collaboration. More importantly, the process involves both individual judgments about the issues as well as a time element — when things need attention. A process that involves a steady flow of predictable information from third-parties and/or significant lag time in response to requests frustrates the ability of actors to control their problem-solving activity. People involved in decision-making need the ability to get answers to their own questions, not just generic ones. When new problems are represented in forms that already exist, the tendency to think in new ways about them is dampened.

With so much discussion about the increasing speed of business and the rapid changes in the business environment, packaging BI doesn't make much sense. Better analytical tools that allow non-technical people to manipulate models, not reports, are clearly being called for. Simon adds:

> *"The very first steps in the problem-solving process are the least understood. What brings (and should bring) problems to the head of the agenda? And when a problem is identified, how can it be represented in a way that facilitates its solution? The way in which problems are represented has much to do with the quality of the solutions that are found."*

First there is the issue of which problems (or opportunities) get attention. The next step is representing the problem in a way it can be understood by others. BI is not arranged this way today. It may reach a wide audience, directly and indirectly, but its use is tiered by roles — power user, analyst, report reader, etc.

"The way in which problems are represented has much to do with the quality of the solutions that are found."

– Herbert Simon

In large organizations, business goals used to be pretty straightforward, and were set at the local level. Now those objectives are often set at the corporate level, and involve trade-offs between risk, resource constraints, opportunity costs and other factors. BI tools are mostly limited to aggregation of individual plans and lack more sophisticated analytical capabilities.

Today's business climate is more highly regulated than ever, requiring compliance with more new regulations, with stricter and more complex rules, shorter deadlines, and with greater consequences for non-compliance. Decision strategies — such as customer management strategies, for example — change more frequently and more rapidly to deal with competitive forces, environmental changes and changes in your customer base. Decisions that were once "owned" by a single group may now be "shared" by multiple departments, and may have to be coordinated across channels and regions. In the past, decisions were handled with manual review processes — now the volume makes that impractical. The "value" of a decision could once be measured in terms of the cost and time needed to make it — now there are other objectives as well.

Best Practice #1: *Recognize that collaborative decision-making may require something different that the traditional deployment of packaged BI. Look for processes and solutions that enable people involved in decision-making to get answers to their own questions, not just pre-determined ones.*

The "value" of a decision

could once be measure

in terms of the cost and

time needed to make it —

now there are other

objectives as well.

EXTERNALIZATION: VOLUME AND DIVERSITY

Today all large organizations are doing business on the Internet. As a result of open standards widely adopted by vendors large organizations, business is transacted with suppliers and customers over the web. But it is not just large organizations anymore and the connections are not just to the partners you know. The ability of almost anyone to connect to anyone else means that new customers can find you, but new competitors can too. All of this activity fuels an explosion in connections, new practices, and data integration challenges.

In the past, these electronic data connections, when they existed, were hosted by proprietary systems and conducted according to rigid formats, such as EDI (electronic data interchange). They were expensive, limited in what they could provide and difficult to maintain and enhance. These relationships formed when one party had clout, such as a large manufacturer or retailer, and could dictate the standard. Other parties found value in joining, but membership was by invitation only. The organization was hierarchical, with information flowing from many points to one, or vice versa. Small organizations were generally excluded.

This affected the amount and quality of data organizations managed. While EDI data was strictly specified, the quality of it was not always good, but the volumes were small because the proprietary networks that carried the messages had very low bandwidth and transmission was slow and expensive. But the volumes that are transmitted now are up by orders of magnitude. Today, anyone can connect to anyone else. Just in the US, there are at least 2.5 million small companies, and most have an online presence.

The ability of almost anyone to connect to anyone… fuels an explosion in connections, new practices and data integration challenges.

Moores' Law

Floor Space Utilization for 1 Tb

Hard Disk Drive Density Trend
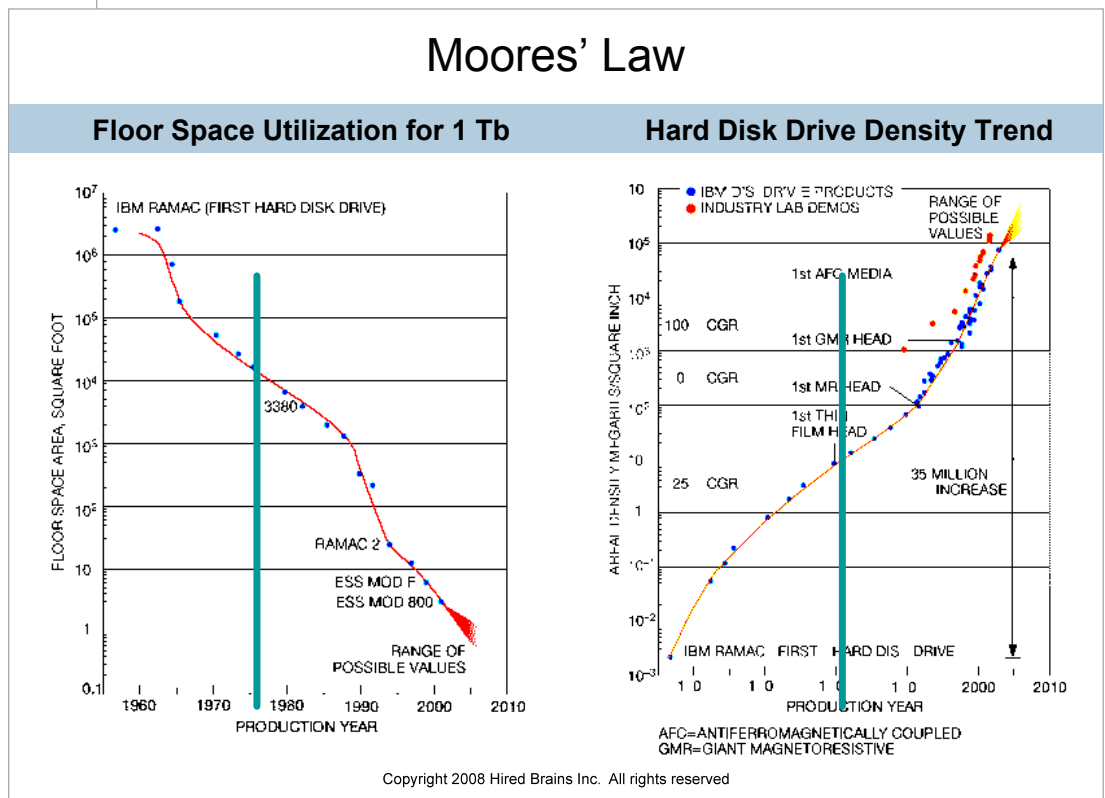
These two charts depict, on the left, the floor space needed to house drives to store one terabyte of data, and on the left, the disk drive density in the corresponding time periods. It starts in 1966, with the invention of the disk drive, until today. The blue line represents the date when data warehousing was first conceived. This may not seem so dramatic for a period of more than 40 years, but when you consider that the vertical axis is in log scale, the enormity of the difference is obvious. In other words, what took 10 million square feet of raised floor, water-cooled data center space can now be squeezed onto a device just 4 inches square. Not only is this trend expected to continue, it is accelerating.

Two important observations can be derived from this. First, we can obviously store a lot more data at much less cost each year. Second, and more important, is that ideas like data warehousing were formed when we were still managing from scarcity — never enough hardware. We no longer are, yet most current data warehouse and Business Intelligence best practices are still based on the idea of preserving physical resources, a concept that is limiting our thinking.

Best Practice #2: When planning for collaborative decision-making BI, question whether your design decisions are still influenced by scarcity. Start with the totality of what is desired and work backwards if necessary. And don't forget to leave some headroom — analytical requirements grow over time.

THE OTHER MOORES' LAW

But having the resources to use information is not enough. Calvin Moores , a pioneering American computer scientist, whose statement about information retrieval became known as "Moores' Law" stated, "An informational retrieval system will tend not to be used whenever it is more painful and troublesome to have the information than not to have it." In fact, there was a corollary to it, too: "Where an information retrieval system tends not to be used, a more capable retrieval system may tend to be used even less."

This is a concept that is never discussed in Information Technology. Having information can be painful and troublesome, like covering your eyes at a horror movie. If you have information, you must first digest it, which is not always easy. You must then try to understand it. To do this, you may have to think about it. The information may require that you make decisions about it or other information. The decisions may require actions in the way of a troublesome program of work, or trips, or painful interviews.

Understanding information may show that your work is wrong, that your boss is wrong, or that your work was needless. If nothing else, information piles up on your desk — unread. It is a nuisance to have it come to you. It is uncomfortable to have to do anything about it. Finally, if you do try to use the information properly, you may be accused of puttering instead of working. Then in the end, the incorporation of the information into the work you do often may not be noticed or appreciated. Work saved is seldom recognized. Work done — even in duplicate — is well paid and rewarded.

This seemingly negative perspective is not all bad. Being successful means recognizing how success is measured. Before embarking on a campaign to "democratize" information in your organization, don't forget Moores' Law, but don't let it dissuade you, either.

Best Practice #3: *Prevailing models of "build it and they will come" will not suffice. Engaging people in your organization is not only possible but necessary.*

Moores cautions us not to be arrogant or cavalier in rolling out information systems to people, which is good advice. However, there is a dilemma here. How are we to deal with this embarrassment of riches afforded us by Moores' Law?

> *"An informational retrieval system will tend not to be used whenever it is more painful and troublesome to have the information than not to have it."*
>
> *– Calvin Moores*

THE NEED FOR DATA VISUALIZATION

As the amount and complexity of data increases along with the speed at which it arrives, canned, static displays, whether they are reports or charts, are not adequate. The only way to understand voluminous and/or complex data is visually. A human's cognitive ability to scan a report or a table and comprehend the whole picture is limited. If for no other reason, visualization can condense the visual field of large datasets and provide presentation that can be scanned and comprehended quickly. In addition, it is much easier to place data of different types in the same display, such as historical and real-time events in a single picture. Patterns can emerge visually that cannot be detected in the detail. Visual displays can also act as a type of aggregation, providing opportunities for navigation, visually, through the data.

But that is only the first step. In order for visualization to work, people who use it have to have the capabilities and skills to understand the information that underlies the display. In addition, information has to be useful for the decision at hand. Visualization has to be interactive. It must respond to interactive navigation. And interactive navigation is not possible within software that does not understand the relationships behind the data, the models.

The only way

to understand

voluminous and/or

complex data

is visually.

FIGURE II:
Integrated visualization and analysis enable not only faster, real-time analysis but also allows analysis of higher quantities of more complex data.

The Need for Data Visualization
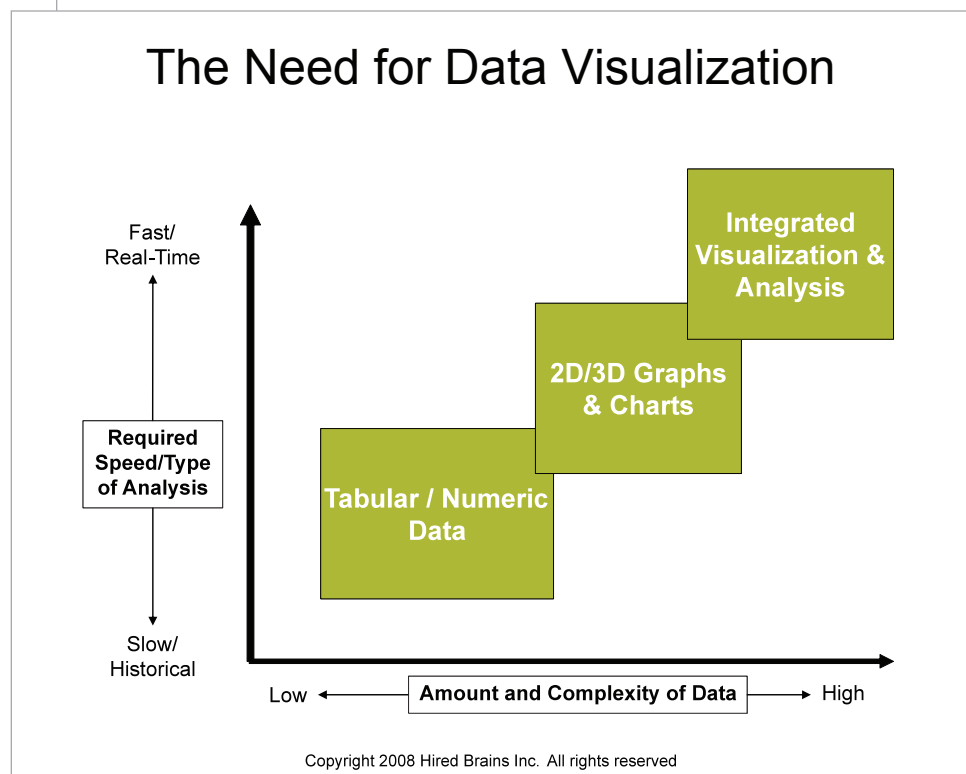
And finally, all the great visual exploration in the world isn't worth a thing if it can't be shared with others.

Best Practice #4: *Voluminous, complex data are best understood visually. Provide interactive visualization as a means of analytics.*
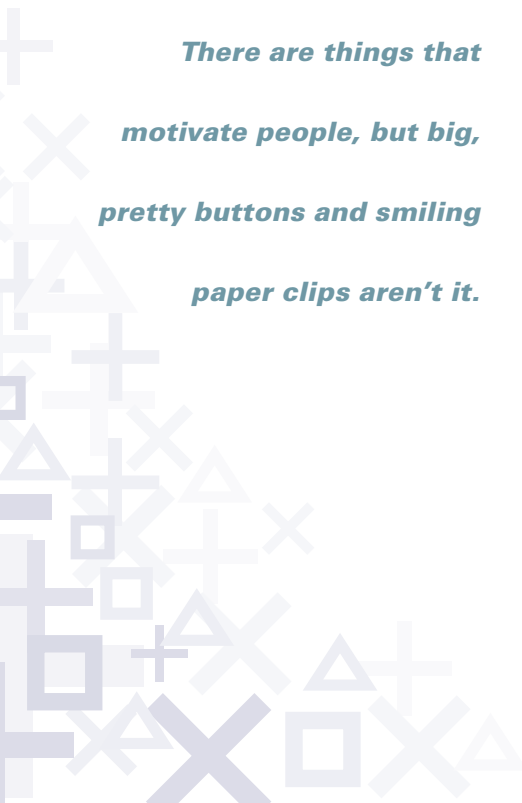
COMPETING ON ANALYTICS

Tom Davenport in his best-selling book "Competing on Analytics," suggests that companies can no longer compete on traditional discriminators like customer service, product development cycle time or even regional semi-monopolies because copy-cat competitors can arise anywhere in the world, quickly, because of technology. For this reason, competing with analytical ability, to understand your business processes, to force out waste and latency and a host of other activities, is now key to success. Analytics is now respectable thanks to him.

So how do you compete on analytics? First of all, you have to redesign analytical methods. Today, they are largely singular, disjointed, personal efforts. They have to become collaborative, iterative, and fast. Out-of-date best practices are so ingrained, that industry organizations are still giving awards to organizations using these old-school methodologies. An award-winning BI recipient at a conference recently claimed that 95% of their queries execute in under five minutes. They have 24,000 queries a day. That means that 1200 queries a day take longer than five minutes. So the question you have to ask is, is there a better way? Are people tolerating a 100 hours a day of latency because they think that's normal?

It's normal because BI and data warehousing are still firmly rooted in batch processes and staged data. The rise of performance management (alternatively named enterprise, corporate and/or business performance management — EPM/CPM/BPM) is a good example. No one can argue that consistent, understood metrics presented in dashboards are a bad thing, but is this the best we can do?

Analytics has historically been done by "quants," those with a mathematical and/or statistical background. In fact, analytics can be performed by anyone, and good analytical software exists that can extend from the novice to super-expert level. No one should build analytical models they can't understand, but that is more of a management issue than a technology issue.

Best Practice #5: *What you think is a best practice today may not be. Reevaluate your analytical processes as well as your expectations.*

Out-of-date best practices are so ingrained, that industry organizations are still giving awards to organizations using these old-school methodologies.

WHAT IS MISSING IS *TOOL ELASTICITY*

People in organizations possess not only a wide range of skills, but also a wide range of skill levels. The resident expert on pricing certain chemicals to the pharmaceutical industry may be a novice at operating a scorecard. An associate in the graphic arts department might have graduate level training in statistics. A collaborative tool for analytics must accommodate that diversity and do it seamlessly. It has to stretch to accommodate not only people with different levels of skill, but also be elastic enough to adapt to their varying level of skill at different tasks. It is impossible for people to share their analysis and results otherwise. It's also impossible to incent people to use analytics in their work when they find the tools imposing or intimidating. This goes far beyond the issues of "ease of use." There are things that motivate people, but big, pretty buttons and smiling paper clips aren't it.

Best Practice #6: "Easy to use" without "useful" is shelfware. Use tools and solutions that stretch to accommodate not only people with different levels of skill, but also be elastic enough to adapt to their varying level of skill at different tasks.

There are things that motivate people, but big, pretty buttons and smiling paper clips aren't it.

WHAT IS BI?

BI, for the most part, pulls data from a data warehouse (or mart), though in many cases BI operates without these formal structures and accesses data from specially prepared structures such as multi-dimensional databases or "cubes," informal extracts and even flat files. One could say that BI in this mode is still in version 1.0.

BI 1.0 Assumptions

Golden Triangle of Business Intelligence (1.0)
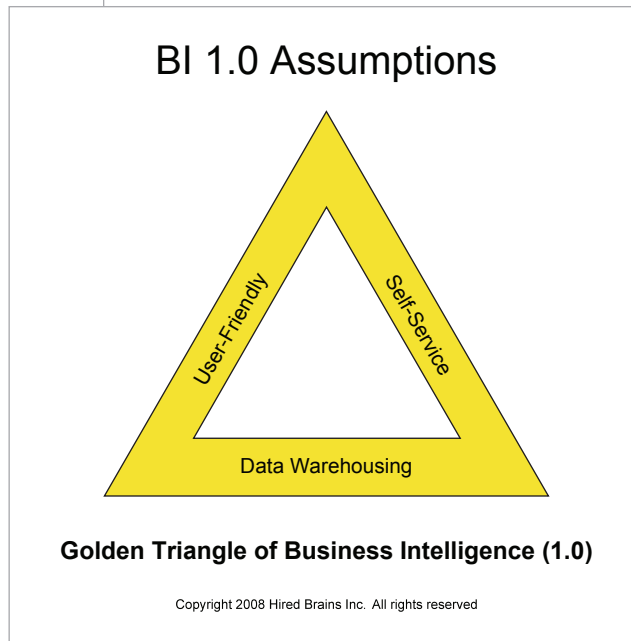
These are the foundations of BI 1.0:

1. Good, persistent data, typically the result of a parallel discipline, data warehousing, takes time and effort and is a complex process performed by information technology (IT) which is meant to be a single repository of integrated data that represents a sort of official version of correct information, often referred to as "the single version of the truth."

2. The idea that the complex data models that underlie such an achievement can be made "user friendly" through aggregation and user interfaces so that those without IT training can understand them and interact with them to satisfy their needs for information directly.

3. That non-technical users have the time and facility to learn how to use these various tools, perform their own exploration and analysis without the need to rely on more technical people and, most importantly of all, that this is sufficient for them to make better decisions that will aid their organization.

Noble goals, but a little behind the times.

BI has been fairly innovative in the past decade, but areas of weakness still remain, such as collaboration (people working together instead of individually), workflow (making analytics part of the wider set of activities) and tool elasticity which we just discussed.
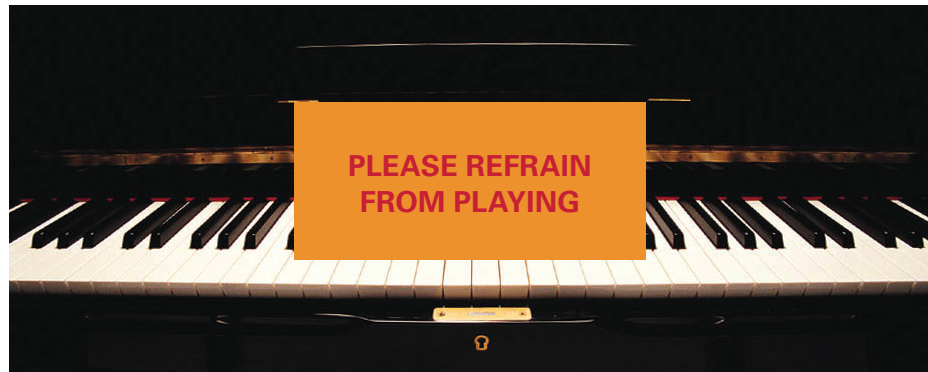
With all of the time, attention and money spent on BI, it is reasonable to ask, how is it doing?

HOW IS BI DOING?

Not so well.

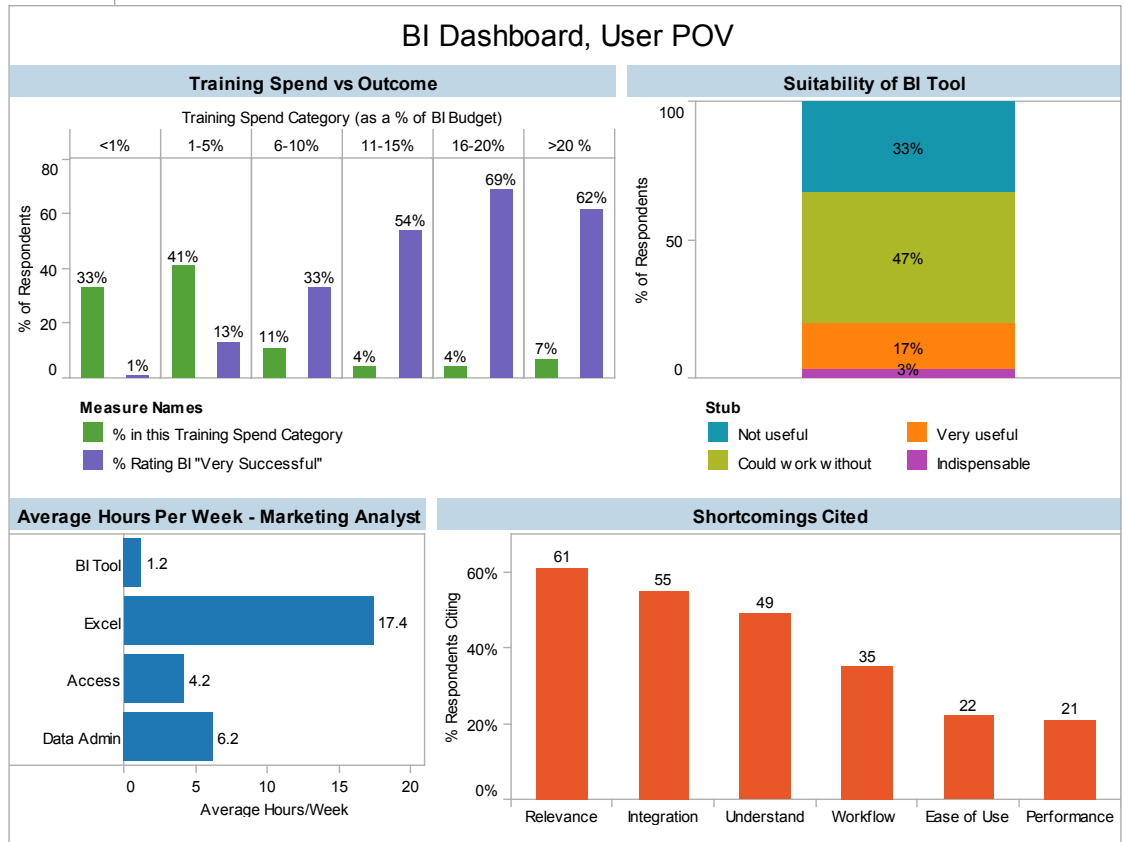User Perception of BI

PLEASE REFRAIN FROM PLAYING

Fewer than 20% of the identified "knowledge workers" in an organization use BI. Some believe the number is lower than that — 20% is derived from the number of licenses sold, but many, including the OLAP Report, believe that BI shelfware, software purchased but not used, may be as high as 50%.

Once we get beyond the "knowledge workers" the impact of BI is non-existent. There is a lot of energy being generated about "operational BI," where BI becomes embedded in operational systems or, alternatively, non-knowledge workers use BI in their work. These are great ideas, but to make them happen will require some elasticity that isn't there now and, even more importantly, an effort on the part of the vendors and the IT organizations to learn a lot more about the work people do, rather than relying on some simple models derived from some conceptions that may not be true.

Consider the following study:

BI Dashboard, User POV

Training Spend vs Outcome

Training Spend Category (as a % of BI Budget)

<1% | 1-5% | 6-10% | 11-15% | 16-20% | >20 %

% of Respondents

33% 1% 41% 13% 11% 33% 4% 54% 4% 69% 7% 62%

Measure Names
- % in this Training Spend Category
- % Rating BI "Very Successful"

Suitability of BI Tool

% of Respondents

33%
47%
17%
3%

Stub
- Not useful
- Could work without
- Very useful
- Indispensable

Average Hours Per Week - Marketing Analyst

BI Tool 1.2
Excel 17.4
Access 4.2
Data Admin 6.2

Average Hours/Week

Shortcomings Cited

% Respondents Citing

Relevance 61
Integration 55
Understand 49
Workflow 35
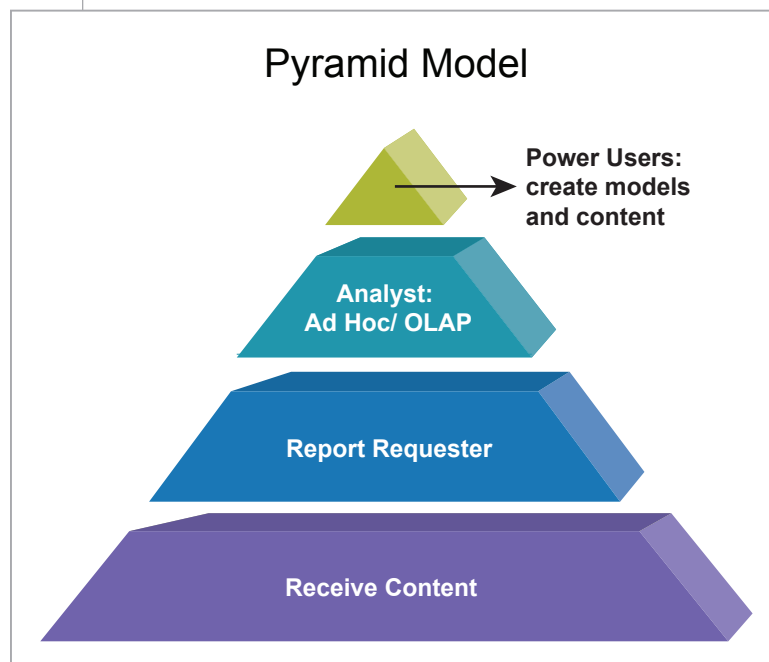Ease of Use 22
Performance 21

In a survey conducted by Hired Brains in 2004, some illuminating sentiments were revealed about the state of BI in large commercial organizations in the USA. Though this data was gathered four years ago, it uncovered some fundamental issues which are still at hand. For example, even in organizations with an enterprise-level BI program in place, on average, analysts spent six to ten times as much time using Excel as the BI tool, for which the average usage was only an hour per week. In addition, only the sliver of the survey population found the BI tools "indispensable," while almost 80% answered "not useful" or "could live without it."

Upon reflection, one might assume that the cause of this pretty much unanimous ennui with BI was the result of the classic problem areas—performance and ease of use. It came as a complete surprise, though, that those two perennial problems areas were the least bothersome. Relevance, integration and understanding were the most often cited shortcomings of BI. Integration was understood by the survey participants to mean integration with the rest of the work they do. In follow-up interviews, it was revealed that lack of understanding of the data, of the underlying data models and the context or semantics of data provided were barriers to use. But relevance was the most often cited issue and it meant, simply, that BI as delivered did not solve any problems.

The lack of relevance can be traced to the gap between IT's focus on operational systems and business people's need for information for decision-making. Bridging this gap will take the combined efforts and cooperation of everyone but, more than anything else, will require the recognition by technology providers, practitioners and senior management in companies, that analytics has to be institutionalized in organizations, not just delivered. That means that project plans for BI have to make provision for extensive change management, and systems integrators have to have both the skills and the portfolio to make those changes happen.

Unlike a new General Ledger system that cuts over into production at a certain date, a BI system can be undermined by lack of interest and participation, with users staying with older, less efficient or less accurate processes, especially spreadsheet and personal databases. Part of the problem is the way the BI industry views people, the work they do, and how to segment them.

Pyramid Model

Power Users: create models and content

Analyst: Ad Hoc/ OLAP

Report Requester

Receive Content

The pyramid model of BI is completely inadequate for today's world of externalized business, computer-savvy workforce and constant communication. The concept of hierarchical decision-making, as well as solitary decision-making in most cases, is simply not tenable. Problem-solving and decision-making happen at every level of today's flattened and distributed organizations. The second word in the phrase "business intelligence" is, after all, intelligence. What does it mean to provide "intelligence" to people and operations? How do systems become intelligent? The enemy of intelligent systems and organization is stasis. Becoming intelligent involves collaboration, sharing, and the ability to publish and modify analytical applications, not just data.

All four steps of the decision process elaborated by Herbert Simon and many others can be satisfied with analytics software that allows people at any level of skill to perform the operations they need to. Super-users will always have a role and most people in an organization will never develop skill or interest in pursuits like stochastic processes or simulation, but framing a problem and building a model can be a very simple process. Today, most BI efforts are driven by data, not by models, and the BI tools user interfaces, best practices and training are aligned with this approach. Toppling the pyramid means breaking through the data-only model and finding ways to distribute models and applications that can be used by everyone.

Best Practice #7: *Successful collaborative BI is more than the right software rolled out to certain segments of users. Count on needing to manage change, and check that your organization is committed to collaborative analytics.*

And there is another roiling issue that needs to be addressed — generational change.

There is another roiling issue that needs to be addressed — generational change.

ENTER THE MILLENNIAL'S

The attitudes of the NextGen vs LastGen towards software in the workplace, particularly enterprise software and BI, are very different. NextGeners are comfortable with technology and expect it to enrich their experience, remove drudgery and handoff's and to enhance their ability to be creative and effective. They are quite effective with software, but also very critical. You can expect in the future that these workers will prefer "mashups," the ability to pull the functions of various things at hand to solve a problem, over trying to solve everything with one favored tool.

Figure VII: Today's younger workers have completely different attitudes and workstyles with respect to technology.

NextGen vs. LastGen

NextGen	LastGen
Raised on technology	Suspicious of it
Filling the cubicles	Fleeing the cubicles
Not afraid of software, but very critical	Resist software solutions, selective
Know how to sew things together	Look for all-in-one solutions
Seek a technology solution first	Seek a 'go-to-guy' first

The previous generation dug in its heels over technology. While some innovations were widely adopted, such as spreadsheets and email, almost every other type of business oriented software was met with resistance, as in the case of BI and especially analytics, at least 80-90% of knowledge workers, and perhaps more, never developed a facility for using the tools except in limited ways, such as exporting data to their spreadsheets.

Likewise, those who grew up with technology seek a technology solution rather than relying on the assistance of a handful of experts who have taken the time to learn the nuance of tool.

Best Practice #8: *Pay attention to the workstyles of the younger generation of workers. If you can engage them with your solution, you probably have something approaching collaborative BI.*

CONCLUSION

The old saying is that first impressions count. In "Blink," Malcolm Gladwell illustrated how first impressions often need some care and feeding and makes the distinction between how people initially react to something and how they may ultimately feel about it. Initial reactions to the breakthrough television shows The Mary Tyler Moore Show and All in the Family were very negative. In the long run, people didn't hate the shows, they were just stunned by how different they were. The conclusion is that first impressions shouldn't be taken at face value — they need interpretation. But this is the weakness of technology deployments in organizations, especially in the field of BI and analytics where adoption can be seen as somewhat optional. After the initial rollout and gratuitous training, people are left to their own devices and first impressions. Clearly, a program to move people past first impressions to a more reality-based assessment of the utility of analytics is needed.

Breaking through the BI pyramid is impossible by merely suggesting it's the wrong approach. The solution is to provide the right approach and allow people in organizations to finally be able to do the work that they've been told they should do — act independently and collaboratively, move with swiftness by being informed and leverage the wealth of technology that is available today to assist them. Technology and service providers must re-educate themselves in the realities of problem-solving and decision-making and start to deal with the situation as it really is, not as their current tools and approaches presume it to be. That requires jettisoning the complex, layered architectures of their products and methodologies and allowing knowledge workers to finally operate at the level that they are capable of.

Best Practice #9: Don't presume that BI's traditional complex, layered architectures and methodologies can be adapted to deliver collaborative BI. Getting more than just BI specialists and hard-core analysts to use fast, lightweight, collaborative BI means recognizing how users really operate and finding solutions to fit that reality.

The solution is to provide the right approach and allow people to finally do the work that they've been told they should do — act independently and collaboratively, move with swiftness by being informed and leverage the wealth of technology that is available today to assist them.

ABOUT THE AUTHOR

Neil Raden is the co-founder of Smart (enough) Systems LLC, a research, advisory and consulting firm. In the past he's been an actuary, a software developer, a DSS consultant, and the founder of Hired Brains, a systems integration company specializing in data warehousing and advanced analytics. He has personally designed dozens of data warehouses and large analytical applications in a broad range of industries and functions. He is a frequent speaker, an active consultant and a widely published author. His current book, "Smart (Enough) Systems: How to Deliver Competitive Advantage by Automating Hidden Decisions," co-authored with James Taylor, was released by Prentice Hall in June, 2007. He can be reached at *neil@smartenoughsystems.com*.