



---

Guided and Open-ended Analytics:  
Serving the Real Users of  
Business Intelligence

**AUTHOR:**  
Neil Raden, Hired Brains, Inc.  
**DATE:**  
September 2009

**“We need an easy-to-learn, data visualization application that (can) deliver key metrics and KPI’s to virtually all our offices without requiring high levels of IT support.”**

***Ian Campbell,***

***Sr. Manager – Growth,***

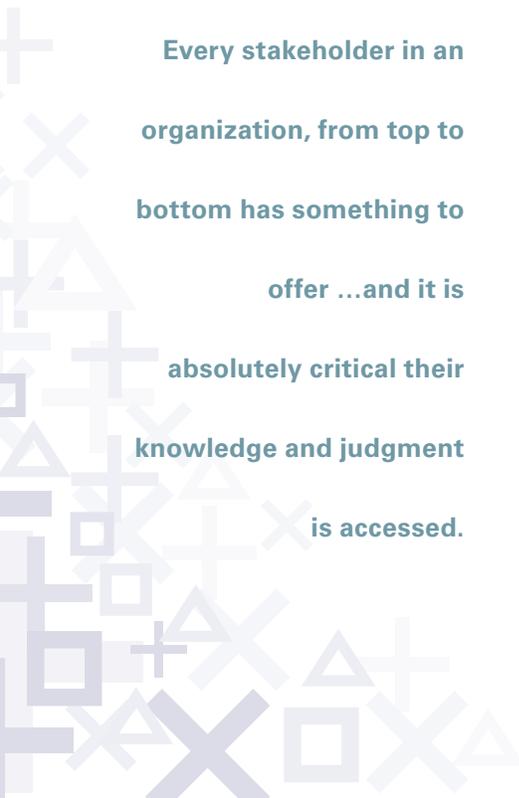
***Ray White.***

## **EXECUTIVE SUMMARY**

It may be a cliché that change is a normal part of business, or that the rate of change is increasing, but tools and techniques for developing strategies to cope with this level of change are being constantly devised and enhanced, and they need to be. Many existing reporting systems in organizations are not only based on a static view of the business, they are architecturally rigid, utilization rates are low and rather than an effective means to cope with change they are a major impediment to it. At certain levels, consistency and repeatability are not only desirable; they are often mandated, such as regulatory and statutory reporting requirements. But reporting at the operational level, reporting to inform decisions and reporting to formulate and monitor strategies must be capable of rapid adaptation. In short, for a reporting architecture to be useful, it must exhibit the features of both fixed, consistent reporting as well as on-the-fly improvisation. This is, as the wise man said, easier said than done.

Measuring performance, investigating variances, looking over the horizon and making plans accordingly are fundamental activities in a knowledge-based economy. The starting point is acquiring a deeper understanding of the causes and effects of things, both internal and external, and having it spread throughout the organization. To do this requires, among other things, tools that provide the promise of self service to evaluate, investigate and share. However, not everyone is capable or is interested in building models or maintaining analytical applications they’ve developed.

Getting the job done, then, requires a mix of tools and approaches. For those who have an analytical perspective but not a technical one, some sort of guided analytics through data and models is called for. What are currently referred to as “dashboards” represent a good compromise of function and aesthetics. A smaller constituency desires to not only create analyses, but to share them with those who are not so inclined. This sort of open-ended analytics allows an analyst (or really anyone with the understanding of the data and relationships in the organization) to author their own analytical scenarios for themselves or for sharing with others.



**Every stakeholder in an organization, from top to bottom has something to offer ...and it is absolutely critical their knowledge and judgment is accessed.**

The focus of this paper is to:

- describe the use of these tools from a cognitive point of view,
- examine the differences between guided analytics and open-ended analytics and
- propose some “best practices” for deployment and use.

## **THE ANALYTICS MANDATE**

In 2009, in the midst of a worldwide financial crisis still unfolding, its effects still emerging, organizations are taking a closer look at their operations than ever before. Clearly, looking at cost structures with the aim to trim wherever possible is usually the first order of business. However, with the realization that uncertain economic circumstances can also provide opportunities to look for areas of growth and advantage, proactive organizations seek ways to get a better understanding of their markets, their customers and their suppliers. Over and over again, it’s been proven that the most successful organizations excel not by framing the perfect strategy and executing it flawlessly, but by placing themselves, using whatever means, in a position where they see and exploit attractive opportunities before the competition. This is only possible when those opportunities are visible, and the viewers know what they are seeing. That requires a great deal of monitoring and much of it is situational and real-time. There isn’t time to take six months to model and implement the information or to adjust.

Even those organizations that do not operate under an intensive competitive model, opportunities to advance the prospects for performance of the organization arise. Being able to see those opportunities, and especially the events that act as leading indicators or early warning, vastly increases the odds that the strategies will be put in place to take advantage of the situation. Our research has shown that this is most likely to occur when a wide cross-section of stakeholders has access to timely information, especially through the visualization of information, active alerts and a means for collaboration based on actionable business information.

Unlike the last recession, where the availability of detailed information was lacking and assumptions were made from fairly high-level, aggregated information, today there is a nearly limitless

amount data to examine and an abundance of computing resources to do it. Every stakeholder in an organization, from top to bottom has something to offer and with this abundance of analytical riches, it is absolutely critical their knowledge and judgment is accessed. This requires more than a central planning group writing database queries and statistical models.

When living in a time of economic uncertainty, relying on the same set of tools and analyses that were employed before the current upheaval doesn't make much sense. In fact, even without major upheaval, circumstances change at both predictable and unanticipated rates all the time, so it stands to reason that analysis and measurement systems need a certain degree of agility to deal with contingency. The problem is that until now, providing knowledge workers with tools that were sufficiently agile required a degree of skill and practice to use that was beyond most people's capabilities and/or interest. This presents as a difficult optimization problem—how much agility is needed, balanced by ease of learning and ease of use? This is a problem that software vendors in the BI industry have been struggling with for decades.

It is a fact of human nature that some people can create things from a conceptualization and others are more effective using the creations of others as a starting point (after all, even Isaac Newton referred to his monumental work as “standing on the shoulders of giants”). It is no reflection on their intelligence or motivation, it is probably just a combination of experience and brain architecture, and it is a good thing that humans exhibit such variation, or we would live in a very boring world. Based on our research over the years, it is better to celebrate the differences in people than to try to standardize everyone on a single approach. The most successful endeavors have always exhibited this tendency.

## **GUIDED AND OPEN-ENDED ANALYTICS**

In terms of analysis and performance management, we can create a duality, for the purposes of discussion, between guided analytics and open-ended analytics. Guided analytics is defined as cases where knowledge workers can pursue their paths of investigation, create their own persistent templates and even share their findings with others, but the underlying data models, and the semantics of those models, are created and maintained by others. In the case of dashboards, sample designs may be provided for different pre-defined areas and people can either use them as is, or modify and

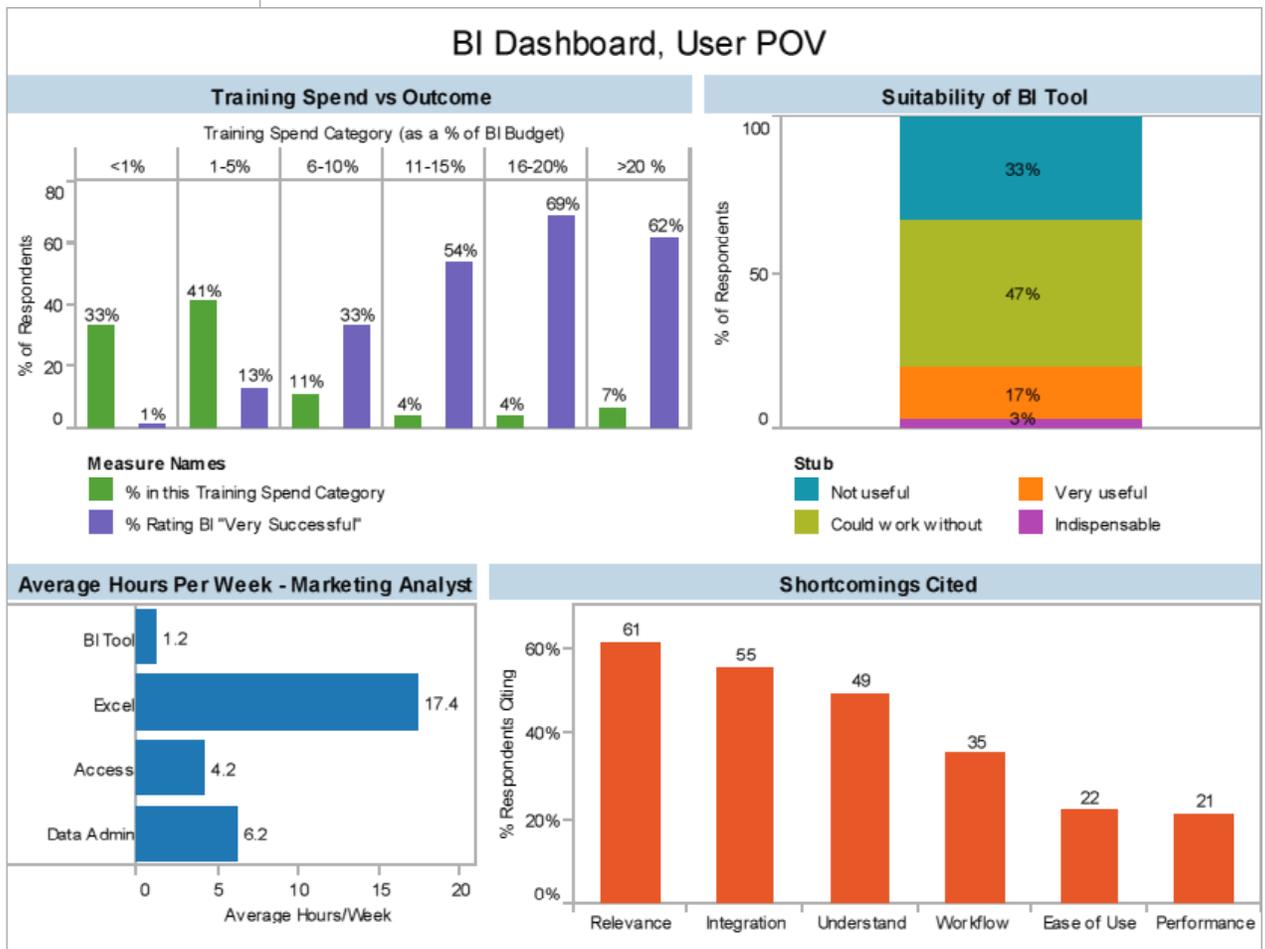
enhance as they wish, but the process has to be intuitive, instantaneous and forgiving.

Open-ended analytics is an extremely useful process for those who have the temperament for it. It generally does not and should not involve “programming” though at certain integration points, most BI platforms require it, either SQL for data issues or Javascript or other web-related languages for creating distributed objects. With open-ended analytics, the knowledge worker is required to understand the semantics of the various models and construct their own analyses. This can range from modifying the templates of others to creating models and metrics from scratch. Though open-ended analytics allows the analyst to explore and synthesize, the process is seen as either too time-consuming or too daunting for many of those who need access to good synthesized information.

## **DRAWBACKS IN OPEN-ENDED ANALYTICS TO-DATE**

Business Intelligence tools, and OLAP in particular, were targeted at filling this need for user-driven BI, but after fifteen years, it is clear that the directed use of these tools (as opposed to passive viewing) has not permeated organizations beyond the 10% or so of knowledge workers who have the time, temperament and desire to master them. While analysts report that this number is growing, our research indicates that the use of directed BI is stagnant and that the perceived ramp-up in usage from the incumbent BI vendors is due to licensing of newer and/or acquired products that are more passive in nature. We also measured five major areas where BI tools were most often found to be lacking for knowledge workers: relevance, integration, understanding, performance (relative) and ease of use.

One thing that BI has never had to deal with was “mission critical” service level requirements. Because BI was typically conducted offline from operational systems, the only two performance criteria that were measured were cycle time to refresh the data warehouses, data marts and other downstream data containers, and query time. Cycle times had to fit into an update window, typically overnight, even for weekly or monthly data, and numerous approaches were adopted, from hardware upgrades to staged rollouts of various



**FIGURE:** Taken from one survey of over 600 analysts, managers and decision-makers, this dashboard shows evidence of one theme – that traditional BI is not serving them well at all. A whopping 80% say their BI tools are not useful or that they could work without them..

levels of aggregated and partitioned data, indexes, etc. Query response times were usually benchmarked against what was possible before, therefore the constraints were loose. In time, however, query response greater than a few minutes became an issue, but the more complicated analyses were still free to run considerably longer. To the extent that analysis can be conducted in the somewhat leisurely manner it always has, performance requirements for next generation BI will not change. But for all of the new applications of BI, such as process measurement and management, unattended decisioning and analytical and data services in an enterprise architecture, the excuses of old ("this was not even possible before") no longer hold up. BI engines have to provide blistering performance.

If a BI engine invokes another service, such as a relational database, to perform some functions, it must be able to request those services in the most efficient manner and it must handle the

queues and the results in a way that adds as little overhead as possible. BI as part of tactical, operational processes does not have the protection and lower performance expectations of classic BI. Next generation BI technology will require a great deal more attention to this performance issue and most BI tool vendors will need to add this competency to their development teams.

From the beginning, technology implementations have been plagued by problems with performance and ease-of-use. Based on our current research conducted at Hired Brains, however, these issues have receded in importance. Whether it is a factor of better software products and more powerful resources, or the application of better technique as practitioners became more sensitive to these issues, is not known for certain. What is known is that the proliferation of BI in recent years increased the population of users and the changed demographics of this community caused a shift in the balance of concerns to three new areas: relevance, integration and understandability.

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 their 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 means there is a clear need for purely do-it-yourself analytics.

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.

## **EXISTING CANDIDATES FOR OPEN-ENDED ANALYTICS**

Because first-generation BI tools were essentially desktop tools and limited to only very small volumes of data, it was assumed that BI was largely the province of summarized data, but it turned out that this was a classic case of the tail wagging the dog. Most business people are interested in answers to questions that

**Most business people are interested in answers to questions that appear, on the surface, fairly mundane, but that can only be satisfied by taking a slice of very UN-summarized data..**

appear, on the surface, fairly mundane, but that can only be satisfied by taking a slice of very UN-summarized data. These questions can actually be quite complex queries against terabytes of data. If left to their own thought processes and not limited by conventional wisdom (from their IT counterparts), more of these questions would be asked, but the limitations of most BI tools have had the effect of constraining people's inquiries. For example, the question, "How many people bought satellite radio in our cars this year," is easily answered with a classic, summarized BI approach. But if this question were posed by a real product manager, it might be phrased more like, "Of the purchased satellite radio options, rank them by the top five other high-end options purchased with them and the other typically purchased options forgone, and dampen the effect of the spring promotion."

The design paradigms of most BI vendors date back 15 years or more. That isn't to say that the code bases of their products haven't been drastically or even completely rewritten since then, but upon close inspection, there is still the distant echo of computer architectures of long ago. For instance, 15 years ago, the "network" usually referred to the proprietary mainframe network that connected terminals to controllers and on to the mainframe. Using BI required "connectivity" to this environment and the process of extracting data for client-server or PC usage was slow, complex and expensive. Today, the network is everywhere and the computer resources at our disposal to access and publish information are almost limitless. In addition, standards exist to allow this flow of information to happen from diverse sources and targets.

The entire data warehouse concept is still stubbornly based on periodic bulk updates of pre-determined cleansed information that always seems to be one subject area behind the requirements. The data warehouse, and the BI tools that are firmly connected to its processes, are still very useful, and will not disappear anytime soon, but their rigidity and centralized focus are not completely adequate.

Existing BI implementations are only slowly evolving from the old connectivity paradigm to one of instantaneous, universal information flow. To the extent incumbent vendors have provided these capabilities, they are often separate products that do not fully integrate with the existing tools, though there are some exceptions. In addition, the provenance of most BI vendors is firmly rooted in the sales/marketing and finance operations of large organizations. To some extent, the prevailing logical models of BI are still multi-dimensional, hierarchical constructs with weak functionality for time-related functions (such as business process steps, connected events, etc.), or for the sort of flatter information common to transactional systems, streaming sensor data or semi-structured data like product or customer lists. Workarounds have been developed over the years, but in general, these tools are still a handful when stepping out of the standard models.

...a great deal of the BI usage in organizations is driven by the need to populate spreadsheets, not replace them.

To cover their needs, information workers have historically filled in the gaps with personal tools, especially spreadsheets which, despite their broad functional reach, are problematic in a number of ways for guided analytics. Nevertheless, spreadsheets do figure in an environment of guided analytics combined with open analysis.

### **HURDLES TO DEPLOYMENT: SPREADSHEET MANIA**

Of the 150 million business users of Microsoft Excel worldwide, a large proportion are devoted to entering data by hand, extracting data manually from other systems, including data warehouses and even BI reports and performing as report servers. In fact, a great deal of the BI usage in organizations is driven by the need to populate spreadsheets, not replace them. The number one asked-for feature in BI tool evaluations is, “Can it export data to my spreadsheet?” This begs the question, has centrally managed BI been misguided, is there some essential characteristic about spreadsheets and personal databases that is too important to overlook, or are the efforts to convert people from disconnected personal tools a mistake? Enduring and expanding spreadsheet applications drain value from multi-year, multimillion-dollar efforts to develop a data warehouse and establish a standard BI environment. Or, is there a basic mismatch in the solution to the problem?

Usage modes of spreadsheets are not evenly distributed, either. The most advanced features of spreadsheets, such as pivot tables, require set up by the most skilled users. This creates a sort of hierarchy where many people depend on the skills of one person to develop these advanced features. The result is a faux IT environment where the work is divided between technical people and everyone else. This has HR implications, as well: if the person leaves, how can they be replaced? On the other hand, being the expert can act like an anchor, holding someone in a position long after it is time to move them on. Today’s regulatory environment demands more transparent and controlled processes, and spreadsheets are a weak link in the chain of accountability. People without solid training in software engineering and architecture, data integration and data quality, and overall project management are likely to make costly errors of design or omission. The risks are not

limited to excessive cost, but potentially expose companies to loss of customers, disastrous decisions, fines and even potentially a loss of license to operate.

The problem is not limited to spreadsheets. Personal databases play a role too, but spreadsheets are the greatest part of the problem by a wide margin. When personal software is stretched beyond its intended purpose and deployed as a shared application, problems arise that were not anticipated in the design of the software. On the other hand, shared or enterprise software solutions rarely have the fit that a custom-made spreadsheet does. The user interfaces are rarely as intuitive and familiar and the care, custody and control of a centralized IT department can often be interpreted as arrogance and detachment. But only the centralized IT processes can deliver the security, scalability, reliability and consistency that are needed. It is nearly impossible to enforce these attributes on personal applications.

## **GOOD INTENTIONS, NO BIG PICTURE**

When someone creates a spreadsheet for their personal use, or to share with others, they employ both subject matter expertise and elements of design. Whether they actually possess the needed subject matter knowledge to do the job effectively is highly variable. Possessing an understanding of good design for reuse, maintainability, security, endurance and a host of other desirable qualities in systems design, however, is very unlikely. There are those in “Shadow IT” organizations who are, in fact, quite skilled at building durable systems from personal tools, but these exceptional people are the exception to the rule. Eliel Saarinen, a groundbreaking architect of the early twentieth century (and also the father of Eero Saarinen, the designer of the St. Louis Arch, among other things), had this to say about good design:

*Always design a thing by considering it in its next larger context – a chair in a room, a room in a house, a house in an environment, an environment in a city plan.*

Saarinen never saw a spreadsheet, his domain was bricks and mortar, but he hit the nail on the head, so to speak. A perfectly functioning spreadsheet on a personal level exhibits undesirable tendencies, emergent failings, when the model expands into a few dimensions, notably time and activities. To paraphrase Saarinen, one might construct a spreadsheet to exist at the moment it is needed, but the “larger context” it exists in is time, and as things change in an organization, as they inevitably do, the rigid

arrangement of things in a spreadsheet, a real advantage for capturing diverse relationships at a point in time, becomes a barrier that can only be crossed with additional development and/or remediation effort.

The problem is magnified by the proliferation of dependent spreadsheets that are modified and calculated in serial form. Lacking a professional system development methodology or even version control scheme, the whole process is manual and the quality of the effort is wholly dependent on the participants' memory of which sheets are related and where they are, and which are the most current versions.

The future of spreadsheets, however, is bright. In addition to their intended function as personal productivity tools, software vendors, including Microsoft, are moving forward with spreadsheets as integrated parts of business and decision systems. But for the time being, they are not suitable as guided analytics systems.

## **WHAT MAKES A GREAT GUIDED ANALYTICS SYSTEM?**

Again the purpose of this article is to demonstrate the need for both guided analytics and open-ended analytics, not to exhaustively define the former category. At its most basic, guided analytics should let you see what's going on and have an easy way to investigate further:

- Without having to learn a new piece of complicated software
- Without a detailed understanding of underlying data models
- Without having to connect to data sources manually
- With the timeliness demanded
- And to set and receive alerts
- With the assistance of the best possible visualization tools

Guided analytics tools are a little like fish dishes: they may be simple or elaborate, but they are always best when they are fresh. Beyond that, there some characteristics that make guided analytics tools effective:

<b>Attractive</b>	Regardless of the medium for presentation, a guided analytics tool has to convey information in a way that facilitates understanding. Layout, color, graphics, sensible sub-totalling and banding, all of these are necessary. Unnecessary clutter, too much shading or too little – there is an aesthetic that is difficult to prescribe but easy to identify as met or unmet.
<b>Informative</b>	In the same way that charts can waste a lot of ink conveying what could be shown in a table with four values, dashboards can fail to inform by either obscuring the pertinent information in a blur of detail or wasting paper (or its electronic variant) with too much display and too little information.
<b>Reliable</b>	Particularly with alerts that inform only when certain events occur, the notified parties will come to rely on the alerts and modify their normal routines. A missed alert can cause lack of confidence in the system or have even more serious consequences.
<b>Intuitive</b>	Did you need a training class for Google? In fact, how many of your employees even took a training class in Excel? Unlike BI tools, where the user population is more heavily weighted with dedicated users who learn to master its high level of interaction and navigation, guided analytics are designed to inform a user population without a lot of facility for the technology. Any degree of training needed to operate them detracts from their essential nature. Likewise, vendor training is not a major selling point.
<b>Timely</b>	There are two aspects to timeliness, arrival and freshness. A dashboard is composed of its content as well as its delivery and late-arriving information has diminished value. The content itself has to be timely as well. A detail of O-negative blood on hand as of 2 o'clock yesterday isn't much use at 2 o'clock today, or the price of an option that closes at 4PM received at 4:10PM.

## GUIDED ANALYTICS IN CONCERT WITH OPEN-ENDED ANALYTICS

By putting powerful analytical capabilities in the hands of people who have until now resisted using the tools of open-ended analytics, the need for need for specialists, power users and data-czars to provide reports, metrics and templates is greatly reduced. In addition, the burden of maintenance is more easily controlled through a reduction of one-off and custom applications developed in the open-ended analytics tools. A well-behaved guided analytics system incorporates its own change management process, aware of provenance and dependencies so that modifications and enhancements do no adversely affect the operation of existing analysis.

Because the guided analytics environment registers and manages models, definitions and data, quality of the analytics can be tested and assured. The free-form nature of open-ended analytics is useful and necessary, but requires more deliberate effort to attain the same results.

*Some other advantages of guided analytics and open-ended analytics operating together in an organization are:*

- *Sharing of information and technique*
- *Moving beyond low-hanging fruit to tackle the difficult problems*
- *Reduced ramp-up time for employees in new positions*
- *Solving end-to-end problems*
- *Integrating analytics with operational systems*

## CONCLUSION

In “Blink,” Malcolm Gladwell makes the distinction between how people initially react to something and how they may ultimately feel about it. Initial reactions to the television shows *The Mary Tyler Moore Show* and *All in the Family* were very negative, but as history reveals, people didn’t hate the shows, they were just stunned by how different they were. They ultimately were the two most popular shows in TV history. 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. They either don’t get or don’t afford themselves the opportunity to warm up to them.

Clearly, a program to move people past first impressions to a more reality-based assessment of the utility of analytics is needed. 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.

## ABOUT THE AUTHOR

*Neil Raden, Founder, Hired Brains, Inc.*



Hired Brains, <http://www.hiredbrains.com>, provides consulting, systems integration and implementation services in Business Intelligence, Data Warehousing and Performance Management for clients worldwide. Hired Brains Research provides consulting, market research and advisory services to the Business Intelligence, Data Warehousing and Decision Support industry. Based in Santa Barbara, CA, Raden is an active consultant and widely published author and speaker on data warehousing, business intelligence and information technology strategy. He is the co-author of "Smart (Enough) Systems," 2007, Prentice Hall. He welcomes your comments at

[nraden@hiredbrains.com](mailto:nraden@hiredbrains.com).

## ABOUT TABLEAU SOFTWARE

Get a free trial copy of Tableau Software to test the power of rapid-fire BI at [www.tableausoftware.com/trial](http://www.tableausoftware.com/trial). Tableau Software, a privately held company based in Seattle WA, provides business intelligence software applications that are fast to install and easy to learn. The power of web-based data visualization and BI enables business people to quickly make discoveries and share insights from all types of databases including large data warehouses.