

BI 2.0

PERVASIVE, INTELLIGENT AND TIMELY

BY NEIL RADEN

JANUARY 2007

ABOUT THE AUTHOR

Neil Raden is the founder of Hired Brains, Inc., www.hiredbrains.com. Hired Brains provides consulting, systems integration and implementation services in Business Intelligence, Decision Automation and Business Process Integration for clients worldwide. Hired Brains Research provides consulting, market research, product marketing and advisory services to the software industry. Based in Santa Barbara, CA, Raden is an active consultant and widely published author and speaker.

Contact: nraden@hiredbrains.com

EXECUTIVE SUMMARY

As the Internet-enabled economy continues to innovate (despite the premature rumors of its death a few short years ago), it is abundantly clear: that our internally focused Information Technology (IT) methodologies are obsolete. For Business-to-Consumer (B2C) and Business-to-Business transactions (B2B), this has been clear for a number of years, but for Business Intelligence, (BI), the message has not completely sunk in. BI software is not adequately addressing the analytical needs of today's private and public sector organizations. Its foundations are rooted in a pre-Internet era of computing concepts and its development has been too slow to keep up with both business and technology requirements. It is too complicated in its implementation, its architecture and its use, it is too expensive to purchase, deploy and maintain and it doesn't scale to the massive volumes and sources of data and dispersed and varied audiences of users that are needed, including automated processes. However, creative and resourceful alternatives are beginning to appear, a sort of BI 2.0 that is addressing these shortcomings.

In the same way that Web 2.0 is an architectural shift in the Web from a mostly static collection of web pages to a dynamic architecture of participation and integration with cognitive agents and Rich Internet Applications (RIA), BI 2.0 will ramp up existing BI. The current offerings are based on a personal, desktop-based metaphor of a one-to-one relationship between an analyst and retrospective view of information with a hoped-for direct link to decisive action, but such instances are rare. BI 2.0 is positioned to be proactive, real-time, operational and integrated with business processes that can extend beyond the firewall as easily as providing simple, personal analytical tools on an as needed basis with a minimal footprint and cost.

Whether technology created today's business climate or business created the favorable conditions for today's technology to emerge is an interesting question, but what matters is that any organization, public or private, is either seeing or will shortly see their most fundamental theories about their business altered by the rush of internet-driven technology. Unlike the first wave almost ten years ago, this time it is based in reality. It is much easier for competitors to copy a product, geographic advantages disappear, price advantage can dry up overnight. For over a century, good companies planned, and stuck to their plan, but the landscape is different now. Behaviors have to change:

- Paying attention to the most minute details and being alerted to the ones that pop out of bounds, instead of looking at the aggregates
- Being able to reorganize in an instant, or even manage from multiple organizational schemes simultaneously
- Embedding analytical processes in the actual operational processes, making your whole organization smarter and freeing knowledgeable people to do more valuable work

BI is a valuable tool in an organization that's been held back by balky architecture, latency and complex data concepts that cannot be grasped by those without data management training. The answer is not dumbing down BI, it's making it smarter.

THE CURRENT STATE

Most Business Intelligence software today rests on an approach to end-user analysis and reporting that predates current technology platforms and paradigms, especially the Internet. Many of the basic principles of BI actually predate the name Business Intelligence itself, which was coined by the Gartner Group in 1989, and those principles are starting to show their age. Nevertheless, the value propositions of the existing BI industry are composed of a set of three original assumptions, a sort of golden triad of BI. They are:

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 organizations

These assumptions seemed reasonable when they were formulated twenty years ago, but today it is clear that the outcome has been less than hoped for. The vendor/analyst/media loop of optimism puts a great deal of positive energy forward, but the truth is that BI has not had the desired impact on most organizations. Instead, the outcome has been that:

- Data warehouses have proven to be very useful at gathering and integrating information, but the latency of their batch-oriented architecture is intolerable with the explosion of communications and the externalization of business, and keeping them apace with agile businesses is difficult. The single version of the truth is too rigid for externalized businesses and too unyielding for rapidly changing conditions and requirements. Metadata standards and practices to add abstraction to a data warehouse, allowing multiple contexts of “truth” to exist, have not materialized.
- Google and other spectacular Consumer Web success stories have proven that simplicity is tantamount. The maximum amount of value delivered with the minimum amount of complexity is the right answer. Very few people have the time or the inclination to learn more than the basics of BI and continue to rely on IT or a handful of “power users” to develop reports and analyses for them, or they simply bypass the whole structure and use spreadsheets at great cost due of inefficiency and poor data quality.
- Aggregation has solved only part of the problem, and the need for operational BI promotes the requirement for access to much greater volumes of much finer-grained data, which presents a scalability problem for many BI tools and designs.
- Decision-making is a decentralized process, no longer the domain of a handful of headquarters executives and managers. BI tools need to be stitched into a collaborative, distributed and, often, real-time, iterative fabric of informing, acting and reporting.
- Most importantly, the very nature of computing has changed from internal, mainframe application and desktop-based metaphors to one of the internet and the Consumer Web, dramatically affecting what everyone expects from interactive information systems today.

In a typical BI implementation, IT is responsible for managing the resources, especially the data, software, hardware and services, such as security and communications. Subject matter experts, such as marketing and finance or specialized areas such as logistics, actuarial, engineering, product design, fraud detection and a host of others, communicate their “requirements” to IT who develop a design and implement a solution, including the selection of appropriate BI tool or tools. This process is so rigid and so full of handoffs and cognitive gaps that it is not a workable solution for today’s on-demand, flexible, global business environment. There are still pockets of functionality where this approach is useful, but there are important, in fact, critical areas, that are underserved and need immediate attention. BI 2.0 is designed for that mission.

FROM BI 1.0 TO BI 2.0

The active part of BI, or Analytics, is still the preserve of a small group of experts in organizations. Despite efforts to decentralize decision-making through flattening of hierarchies and deployment of decision support tools to assist people in the effort, the process of evaluating information critically is not widely dispersed in most organization. Some thought leaders, like Tom Davenport, believe this is actually a desirable situation, and that centralized analytics is the key to the effective use of current technology. Recent industry messages like BI for the Masses or BI Democracy are counterpoints to this point of view, but either way, it highlights the well-documented fact that BI has not reached very far into organizations.

Existing BI solutions are designed primarily for people who can understand data models and have the time to build analyses from them, recall them for future use and provide information for others. In most organizations, that's about 5% of the salaried workforce. Everyone else uses Excel.

In the long run, it is nearly impossible to understand the long-term Total Cost of Ownership of traditional BI because the upfront costs include not only licensing and training, but all of the enabling costs of infrastructure and development of the data warehouse. The maintenance costs to keep these applications running are actually a multiple of the development costs over time. The actual benefits achieved by BI to-date are somewhat suspect given the low adoption rate.

Moving from BI 1.0 to BI 2.0 will be most dramatic in four areas:

1. Reactive to Proactive: BI 1.0 is designed around a multi-step process of extracting data from source systems and carefully integrating it into persistent data stores in a multilevel architecture that is most often a daily process, at best. For intra-day processing, special exceptions are made, but usually with a loss of functionality or integration, or both. While there is provision for event-based or exception-based reporting, the triggers are usually found in the analytical applications and not actually real-time. In other words, at the end of the nightly batch update of the data warehouse or data marts, further analysis may cause exception reporting, but it is not, strictly speaking, in the actual flow of business.

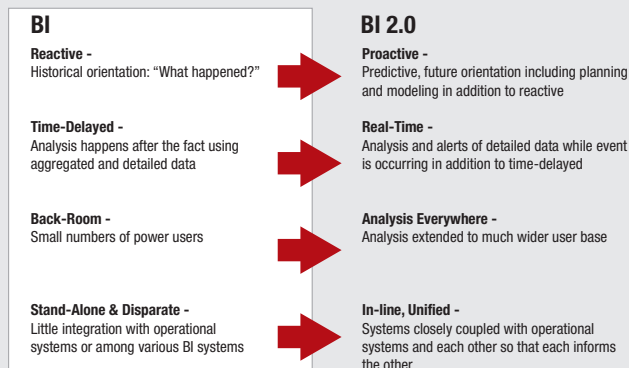
The lion's share of the effort is geared toward historical reporting, most of it passive in the form of pushed or published reports, either static or parameterized at runtime through menus.

BI 2.0 places the emphasis on analytics not architecture. It is capable of reaching out to data anywhere it can be found and performing integration on the fly if necessary. It can leverage not

only data sources, but message queues, logs, web services and many other sources as well. It can leverage query federation, caching schemes or grids. Anyone (or any running process) with permission can simply choose to monitor something in the manner they choose. For instance, someone in purchasing may drag the product codes from two competing suppliers for a given component and monitor the shipments, prices and contractual performances on a dashboard and chose to be alerted by email if there is a problem they define as out of bounds by a certain percentage.

2. Real-Time: Many analytic requirements do not have a real time component. There is often time to consider alternatives, to evaluate what happened or to make alternative plans and seek different opinions. But when decisions can't wait, such as revenue optimization decisions, pricing, cross-sell opportunities, personalization and even many applications that are being discovered everyday with expanding use of the Web, an analytic architecture with built-in latency in a true competitive disadvantage. Operational BI is a good example. Many organizations are finding that their traditional means of differentiation for competitive purposes, such as geographical uniqueness or customer service excellence or product differentiation are no longer sufficient to keep ahead of the competition. Globalization and technology make it much easier for new entrants to copy and compete and staying competitive requires, at a very minimum, keeping the processes humming along at maximum efficiency. Without real-time decision management embedded in the process management systems, that level of efficiency is not possible. BI 2.0, with its standards-based open architecture and powerful analytics is the perfect companion to modern process management tools.

BI 2.0: Extension of Business Intelligence



Source: Hired Brains, Inc. © 2007

Data warehouses can capture massive volumes of highly detailed data, but they provision it to BI 1.0 in aggregated chunks and generally only allow queries at the low levels to a very specialized audience. In a mixed operational/analytical environment like BI 2.0, this scheme will not work. BI 2.0 will provide more realistic models for accessing data anywhere it is needed.

3. Analysis Everywhere: The most powerful features of BI 1.0 were never mastered by more than a handful of “power users” in each organization. The reach of BI 1.0 is fairly wide at perhaps 25% of the knowledge workers, though this is mostly due to the aggressive sales work of the vendors and the depth of work performed is fairly shallow. BI 2.0 will be more pervasive in terms of both depth and breadth because it will present itself in the same way that popular Web applications do – with simple, unobtrusive, and useful interfaces and invisible upgrades. The power of BI 2.0 will reveal itself as people use it, and its user interface will learn how the user works rather than vice-versa.

4. In-Line, Unified: Once BI takes its place as part of the operational/analytical hybrid workflow, time will become compressed and BI will need to inherit the same performance and service levels as operational software. There will be no time to write detailed specifications, model and design a database and go through a development/test/release cycle for every application. Instead, new analyses will be built incrementally, using social networking with your peers, through tagging, semantics and intelligent agents. The ability to grab things on the fly, rearranging your business everyday, even in the middle of a day, will be key competitive advantages.

In BI 1.0, “taking requirements” was a major stumbling block because a solid methodology was never found. Those who were supposed to know the requirements often did not because they could not understand the technology before they could apply it. It was a conundrum. Various attempts at iterative development were proposed, but patience and budgets wore thin. In today’s environment, brilliant success stories like Google, Yahoo, Amazon and EBay arose without requirements, they arose from brilliant original ideas that were derived from emerging technology and standards and were continuously enhanced and improved as the result of competition and user feedback. In BI, every methodology discusses alignment between technical and business stakeholders without any specific guidance about achieving it, but devotes stacks of documentation to gathering requirements from the users prior to a discontinuous set of steps to design and construct a solution. This was a reasonable solution in the application development era, but in the web world, it no longer makes any sense.

ESSENTIAL ELEMENTS OF BI 2.0

BI 2.0 is different from current BI in its economic and licensing model, the way it will be distributed and used in organizations, its more direct role in the operations of organizations, its ability to extend its functionality from the very simple to the very complex and its harmonization with other applications both within and beyond the organization’s boundaries. Users of BI 2.0 will be free of the plumbing of data warehousing and able to concentrate on the work they do rather than learning aspects of data models and databases.

In particular, these elements can be summarized into three categories of Use, Economics and Technology:

USE

BI in the hands of everyone

For BI to be effective its use can’t be limited to a handful of “power users” with distribution of static reports to the rest of the population. Any person, or process or service or even a transaction that needs to be informed by an analytical action has to have access to the BI services at some level. This can either be on an ad hoc basis, a scheduled basis, an exception, a continuous “ticker” or a defined event.

> *BI 2.0 will provide a broad range capabilities tailored to the needs of the audience with the goal of serving every constituency that needs analytics*

Simplicity

The key to success of the Consumer Web, such as Google or EBay, is the simplicity of the offering as a function of the value of the service. Three billion people use Google everyday without having taken a training class. At some level, BI may be complicated, but it does not need to be more complicated than necessary.

> *BI 2.0 will provide simple interfaces based on lessons learned from the Consumer Web*

Closed loop

One assumption in current thinking about BI is that it is sufficient to merely provide information to people. A constant refrain is, “Provide the right information at the right time to the right person so he/she can make better decisions.” Unfortunately, there is a gap between those steps and making decisions. Decision-making is a more complicated process than that. Providing information is not sufficient for making decisions, it is usually a collaborative effort. Even in those cases where one person is the actual decision maker, it is unlikely that the information provided by a BI tool represents all of the information needed to make and implement

a decision. It is usually a more complicated process and BI is typically disconnected from it. BI generally lacks the capability to integrate smoothly with the workflow.

- > *BI 2.0 will integrate smoothly with the workflow so that there is no disconnect between analysis and action*

Operational BI

Operational BI is not preparing reports from data extracted from operational systems. Operational BI is BI embedded, or working cooperatively, with operational systems, often without human intervention. This can take the form of scanning message coming across a queue and watching for things out of bounds, then firing off an alert, or in-line credit authorization, seat assignment, web page personalization or simple underwriting decisions.

- > *Operational BI requires BI systems to perform at transactional speeds without latency, to rationalize their architecture and to free up human resources to do more important work*

Shrink BI interface

Most people in organizations have fairly straightforward requirements from BI and need a very simple but useful interface. Google may be the most widely-used software in the world, but it would have never won an award (prior to its success) for its interface design. Most of these requirements can be service with a lightweight, simple architecture.

- > *BI 2.0 principle – keep it simple*

ECONOMICS

Licensing

The Web has shown that there is an appetite for much lower priced software and services, especially at the front end. Today's high entry costs for BI are an impediment to innovation because, once the investment is made, it is difficult to justify switching. It also poses a serious challenge for the existing vendors to offer BI 2.0 products without cannibalizing their legacy base. Also, at \$1000-\$5000 per seat, BI is currently too expensive and complex to give to everyone in the organization, so its rollout is limited by license cost.

- > *BI 2.0 will follow the Web model of pricing and access will be On Demand*

Speed to deploy

Look to BI 2.0 to be simple, incremental deployments relying on either hosted or Software as a Service offerings, or, if in house, added as services in a standards-based services-oriented architecture. With lightweight components and connectors to existing data sources, the long ramp-up with current BI projects to build a data warehouse first (or modify an existing one) can be avoided and save months or even years

- > *BI 2.0 can be installed and delivering useful output in a few weeks or even days because of its lightweight footprint and reliance on open standards*

TECHNOLOGY

Appliances

One intriguing development is Celequest's BI appliance, an all-in-one box with processors, memory, storage and software that is ready to go to run their entire application. It includes their proprietary software as well as open source operating system, web and application server and relational database. Consider the difference between a vendor supplied box where all of the components are supported, upgrades are handled invisibly and any incompatibility or bugs between the pieces have to be sorted out by the vendor, the ideal "one neck to choke" approach, with a typical data warehouse/BI environment with six or eight different vendors, all on different upgrade schedules and all blaming the other for performance problems.

- > *BI 2.0 appliances are another solution to simplicity and distancing BI implementation from plumbing and technical issues. Continuous, harmonized upgrades allow BI 2.0 to run 24/7.*

Real Time

Real-Time is not necessary for most planning or Performance Management applications, but if making your organization smarter by pushing BI into your operations is on your agenda, then real-time is a key issue. There is no magic in real-time, unless your architecture is designed for batch, and BI 1.0 is a batch world.

- > *Real-time does not require people to stare at a screen 24/7, there are software agents for that, but some of those agents have to think, and they need data. That's where BI 2.0 is applied*

Scale

Without physically moving all of the data needed for analysis and storing it in one place, it is possible to address a much broader range of information. BI 2.0 will still rely on data warehouses, but data warehouses will not longer drive BI. BI 2.0, with its lightweight footprint and ability to distribute its processing, can scale easily and, in some ways, without having to by avoiding many of the steps of current data warehousing/BI practice.

- > *BI 2.0 will be able to leverage the extreme cost advantages of Moore's Law because it is not bound to a particular location or platform*

FEATURE/FUNCTION COMPARISON BI 1.0 TO BI 2.0

DATA:

The most striking change is that BI 2.0 is not subservient to data warehousing methodology and models. BI 2.0 can use data from any source, even data from customers, suppliers, partners and other external sources. Integration of data, to the extent it needed or possible, can be done on the fly using semantic tags and semantic integration engines that are rapidly emerging to support e-commerce. Even the provisioning of data from data warehouses will be driven more dynamic processes based on semantic models of roles rather the rigid, limited user roles in place today. Data gathering agents, operating with “intelligence,” will seek out useful information and learn as they go what is important to an individual or a process.

ANALYTICS:

In BI 1.0, analytics is a personal process where analysts perform their investigations and then struggle to represent their findings to others in yet another application. There is a disconnect between the BI tool and the wider audience that needs to under the results of the study. In BI 2.0, analytical work is collaborative, borrowing heavily from the social networking of Web 2.0, and also includes guides that assist the analyst through the process, making suggestion is a useful but unobtrusive way.

APPLICATIONS:

All of this is provided through the tools of Web 2.0 such as Ajax, resulting in a Rich Internet Application (RIA).

ECONOMICS:

Much lower up front costs, lower TCO through the use of dedicated appliances, open standards, On Demand, Open Source for typical components such as databases operating systems, web and applications servers.

BI 1.0	BI 2.0
ETL; batch; latency	Integration on demand; semantics
Data Warehouses	Dynamic data provisioning based on roles, process, habits and situation
Data Marts	Knowledge gathering agents embedded in distributed islands of data
Personal analytics	Collaborative, guides, closed loop via RIA
Windows clients, “webtops”	Ajax, true web app's
Software license fees, version upgrades	Open source, continuous upgrades, SaaS
Single version of the “truth”	Diverse contexts, dynamically managed Cross-enterprise integration
Analytical apps, dedicated analysts	Ubiquitous analytics
Feature overload	Shift in focus to lightweight components
Queries, Reports, Structured Data	Annotations, derivations, commentary

BI 2.0 WILL CHANGE DATA WAREHOUSING, TOO

BI 2.0 must be oriented towards effective decision-making and analysis, not data-centric concepts like data warehouses. Data warehouses have an important role to play, but it will be in the background, hidden by layers of abstraction. In addition, access to data warehouse data, whether aggregated or mountains of detail, cannot be constrained as the convergence of operational and analytical processing will require both. Data warehouse providers must figure out how to ramp performance of their databases to provide these services at operational levels.

In the classical definition of the data warehouse, data pulled from operational systems was “lightly summarized,” but this cryptic definition was never explained. In practice, it meant it was as summarized as it needed to be to fit into whatever batch time and space constraints there were, but by the time users got their hands on it, it was a lot more than lightly summarized.

Because analytical data is highly summarized and aggregated, the move to operational BI is going to take more than a slight shift in orientation, it will take a whole new architecture and a whole new set of tools and methodologies, not just a trademark. Data warehousing has two drawbacks that have to be addressed. First, it uses lots of resources getting data in, but it can't allow ad hoc query of the detailed data warehouse because it can't perform, then it can't do operational reporting because after the integration of data from multiple sources, it can't tie to the source systems anymore because it lined up with the data warehouse model, so the data was only useful for supporting aggregated data marts.

Raw performance alone is not a complete solution. Hybrid solutions of data warehouses, federation, caching and abstraction are needed to find optimizing solutions to a complex problem.

CONCLUSION

BI 2.0 is based on a few simple principles. To be effective, BI has to focus on simplicity of operation to achieve pervasiveness in the organization and beyond it. The model is the Consumer Web, which provides only the necessary presentation to perform the tasks at hand, and relies on open standards and loosely coupled services to perform the functions, which can be reconfigured dynamically. In the same way the users of the Consumer Web are willing to pay little or nothing directly, the cost of BI 2.0 has to drop drastically from expensive, front-loaded perpetual licenses to pay-as-you-go on demand schemes.

In "Blink"ⁱ, 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. 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.

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 educate themselves on the realities of problem-solving and decision-making and start to deal with 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.

ⁱ Malcolm Gladwell, *Blink*, (New York: Little Brown, 2005)