



December / 2010

Expanding Business Analytics: Supporting ALL Information Workers

Claudia Imhoff, Ph.D.

Table of Contents

| | |
|---|----|
| Introduction..... | 3 |
| Three Categories of Information Workers | 3 |
| Technological Considerations | 8 |
| Deployment Options..... | 12 |
| Summary..... | 13 |

Introduction

Business Intelligence is in its third decade of supplying valuable information to an enterprise's decision makers. The role of Business Intelligence is now expanding to include new areas like predictive analytics giving rise to a broader category of Business Analytics (BA). At first, the users of BA were somewhat limited to the analysts or "quants" in an organization – those information workers with significant technical and data manipulation prowess. They were the actuaries, statisticians, market researchers, and financial analysts, people who were responsible for creating sophisticated algorithms, studying historical information for patterns, determining market trends, etc., and then explaining them to the top executives. We called these information workers, the "power users", and developed BA environments that suited their needs well. Many used mainframe "power tools" to access legacy data sources that were too difficult for the non-technical user.

However, for BA to reach its full potential, its audience has to expand beyond these hard-core users. BA vendors and implementers now realize that everyone in the enterprise is a decision maker at specific points in their work activities. All employees must be considered information workers and need to have appropriate access to the right information to make the best decisions. These information workers are not all the same; in fact, they can be classified into three broad categories – each having specific characteristics and technological requirements. BA environments and the technologies supporting them must accommodate all these information workers.

This paper describes the characteristics of these three categories of information workers and the key BA capabilities each will need. Then the deployment options will be described along with the technological considerations will be detailed to create a truly world-class and all-encompassing enterprise BA environment.

Three Categories of Information Workers

Peter Drucker in his seminal work on the new information age workers

coined the term, “knowledge worker¹”. His definition was:

“One who works primarily with information or one who develops and uses knowledge in the workplace”. He continued, asserting that the knowledge worker was someone who “adds value in the workplace by processing existing information to create new information to define and solve problems.”

IBM expanded on this idea with research into the *collaborative knowledge worker*². Their researchers found that knowledge workers are among a corporation’s top talent, and that ways to improve their effectiveness are needed. These workers rely on the ability to work collaboratively, to leverage relationship capital, and to deliver new solutions from these collaborations. IBM also found that knowledge workers often go outside of “official channels” to find relevant data and create their own informal systems to place the data and analyses in proper sequence.

In my research with Colin White, we have refined these findings further and have determined that information workers fall into three distinct categories – Information Producers, Information Consumers, and Information Collaborators. Obviously, understanding each group’s characteristics as well as their technological needs must be the critical step toward creating BA environments that enable them to perform efficiently. By improving the technologies and understanding the information workers’ work practices, BA implementers can greatly impact the knowledge work component of many jobs.

Information Producers

Information producers are often called the power users in enterprises. These were the original users for most data warehouse and BA environments (highly trained statisticians, analysts, line of business IT professionals, etc.) and they are, in general, happy with those environments. Producers are technologically advanced and quite knowledgeable about data. They have the ability to quickly locate data and can create many types of analytics. They may set up data feeds from an existing BA environment and then create their own specialized environments, using their own hardware, software,

¹ Peter Drucker, *The Age of Discontinuity*, Harper and Row, New York (1969)

² From: “Ethnographic Study of Collaborative Knowledge Work” by Sandra Kogan and Michael Muller, <http://www.entrepreneur.com/tradejournals/article/print/155568052.html>

and BA technologies. The enterprise tends to depend on these individuals to create the information used by other information workers in the organization. Their technical skills allow them to adapt to any business analytics tool though they are reluctant to change once they have an established set of skills.

Information producers tend to be an independent lot; they like being able to “serve themselves” in terms of having access to data, creating analytics, and choosing the appropriate technological environment. They combine their analytic products with business rules and strive to deploy these results in real time to the business. In this way, they put great pressure on the BA environments in terms of performance and scalability.

Information Consumers

Information consumers are task-oriented business users and are more operationally focused. They are becoming a significant user base for BA as more and more companies struggle to bring BA into its operational processes. While information is certainly an enabler for these workers, it is not and probably will never be the sole focus of their jobs. For example, sales representatives will use BA information to determine which of their customers are not purchasing as much from them as in the past. A customer service representative will use BA results to determine the next best product to offer a current caller or perhaps look up their lifetime value score to determine how they should handle that customer's request or problem. Both of these employees will use BA results to help them make a decision but it is unlikely that they would create the analytic to uncover a trend or indicate a cross-sell opportunity; they prefer to have someone else (a producer) create the analytic and deliver the results (the purchasing trend or the next best product suggestion) to them. It is important to note that their lack of technical skills should not be construed as making them less valuable to the enterprise.

Information consumers are often frustrated by the technological barriers they must overcome to simply consume analytic results – they must understand where the information is, learn the tool to access it, find the right report or analytic result, understand what the result means, and then go on with their next operational activity. Often, information producers will create guided decision-making environments for these users to simplify the process.

These BA users, the largest population of information workers, are found at all levels of the enterprise from senior executive to line manager to entry-level customer support personnel. We are just beginning to see real progress in creating BA environments that support these workers – dashboards, intelligence portals, mobile BA are all technological advances that support the information consumer.

Information Collaborators

Many BA vendors and implementers are starting to incorporate some form of collaboration technology within their BA environments. The benefits of combining BA and collaboration are numerous. First and perhaps foremost, the addition of collaborative information to BA analytics makes the results far easier to understand and use by the consumers of information. Just accessing a metric without the explanation of what it means and how it may affect the consumer's processes frustrates and confuses the consumer. Without understanding the context surrounding the metric means that the consumer must make assumptions about that information and that is always a suboptimal situation.

When combining BA and collaboration we need to support a new type of information worker – the *information collaborator*. These information workers are “business innovators” who want to use their experience, expertise, influence, and opinions to add real business knowledge to BA analytics. They are the “commentators” of the business that give not only the history behind a business event but also bring in extra information, even external data, to put it into proper perspective. Often, the collaborator evolves from an information producer or even an information consumer due to their expansive knowledge of certain situations. If available (and easy to use), these important workers will use collaboration technologies to supply the additional knowledge behind the analytics. For example, they could add business tags to analytics, supply links to related and trusted information sources, and bring in quantitative information, even quality metrics, through ratings and expert commentary or opinions.

Basically they add business *context* to the existing information content. Important BA capabilities for these workers include:

1. Relating analytics to a business glossary or business definitions
2. Tying analytics to related information like external sources (e.g., weather reports, social media, competitor information, customer demographic information, key links)
3. Connecting analytics to business processes (e.g., clarifying where in the business process to “inject” the BA analytic before making a decision)
4. Adding commentary to analytics and being able to rate reports, analytics, and even the creators of analytic results.

Since information collaborators are the newest information workers identified as contributing to and needing BA, they are naturally the least supported by current BA technologies. They use both structured and unstructured data, collaboration technology, and are or will be using social media in new and creative ways.

Table 1 contains a synopsis of the characteristics and examples of these three information workers. The scale used for the first two columns is 1 – 5 where 1 is None, 2 is Some, 3 is Moderate, 4 is Good, and 5 is Excellent.

Table 1: Three Types of Information Workers

| Type of Information Worker | Satisfaction with BA Capabilities | Technical Skill / Expertise | BA Activities | Examples of Information Worker |
|----------------------------|-----------------------------------|-----------------------------|--|---|
| Information Producer | 4 – 5 | 4 - 5 | Regression and Predictive Analytics, Experimental Design, Simulation Creation of Other BA Products for Consumers | Business and Financial Analysts, Statisticians, Actuaries |
| Information Consumer | 2 – 3 | 1 – 3 | Basic Drill-down, Simple Queries, Report Viewing | Customer Service Rep, Sales Rep, Store Manager |
| Information Collaborator | 1 - 2 | 3 - 5 | Addition of Social Feedback (Comments, Annotations, Links, Rankings) | Information Producers and Motivated Information Consumers |

It is important to note that the lines between these information workers are blurring – an information producer can become an consumer or collaborator; an information consumer can become a producer or consumer and, likewise, a collaborator can become either a producer or consumer at any given time in their workflow.

So what technological environment can support these three categories of these information workers? As the need for more sophisticated analysis of data becomes ever greater, and as the nature of data itself changes, businesses are rethinking their BI deployment strategy to leverage that data to its full potential and gain a competitive advantage. BI and Business Analytics (BA) solutions are now considered to be critical strategic assets that are essential for business insight: they provide an immediate view of how the business is performing, why it's happening, and what should be done going forward.

Organizations are now reconsidering their strategy from a BA perspective to support new requirements for high performance, availability, reliability and security, the very attributes organizations look for when selecting IBM System z. The rest of this paper will discuss the technological considerations and deployment options available to support the needs of current and future information workers.

Technological Considerations

To expand BA to a much broader audience, the underlying architecture must have certain characteristics and standards allowing it to be easily used, consumed, and understood:

- Common user interface – a shared user interface across tasks makes embedding BA into operational workflows easy and intuitive. The common interface improves the company's ability to meet service level agreements as well.
- Common components and services – reuse of common components like the data warehouse, shared environment administrative utilities (like load utilities), and shared procedures and services all reduce overall costs and skill requirements for delivering BA while they improve security and compliance.

- Common models and metadata – sprinkling diverse, unreconciled data models and metadata throughout a BA environment can cause significant problems leading to errors and misinterpretations in the analytic results. Having shared models and metadata yields control and efficiency in developing new BA applications.
- Aggregated performance information – helps identify performance problems faster and enables more effective problem isolation, optimization, capacity planning, and impact analysis.

Fortunately, today's modern System z environment supports all of these architectural components but let's look at detailed specifics with respect to our information worker requirements:

- Ability to support a variety of business analytics workloads – System z has the right solutions for all categories of information workers, with the capacity to deploy multiple analytic applications on a choice of cost-effective environments. Today's System z mainframe not only can support large workloads on its core processors, but offload a variety of analysis applications to dedicated processors that fully leverage mainframe resources. Information producers who run massive queries must have the horsepower to return these queries within a reasonable timeframe. On the other hand, the short quick queries of consumers, while requiring less horsepower, do require speedy response times. IBM mainframes include an Integrated Facility for Linux (IFL), a processor dedicated to Linux-based workloads capable of supporting the range of BA features for all information workers (reporting, analysis, dash-boarding and real-time monitoring as well as predictive capabilities). In addition, high-speed engines – called the System z Integrated Information Processor (zIIP) and z Application Assist Processor (zAAP) – reduce overall processing costs when data is centralized on the mainframe. The mainframe's Logical Partitions (LPARs) will support an almost unlimited number of workloads run within their own distinct operating environments. These specialty engines dramatically reduce the processing costs for business analytics workloads and data warehousing.
- Co-existence of operational data and BA data – in a mainframe environment, everyday production systems run

alongside a data warehouse. Often, the data need not go through the ETL cycle, but be immediately available for actionable decisions by front-line representatives. Since operational data is captured and housed on the mainframe, this can be made available on a near real-time basis for decision-making or embedded within applications, as well as stored within the data warehouse. Even if some of the source data resides on non-mainframe platforms, exporting it to the System z and performing data integration and clean up there makes the process very cost-effective. Having a single system involved in the majority of data integration, clean up and transformation means the overall process is easier and less costly to manage. In addition, audit and control complexities – along with security costs – are significantly reduced. All these improve the implementation times of BA projects. Rapid time to implementation is the watchword in today's BA environments!

- Capacity on demand – System z allows the dynamic expansion and contraction of resources to adjust for peak business analytics workloads, as they are needed. The mainframe permits the full range of system resources to be committed in a productive manner, rather than withholding processing power “just in case.”
- High performance – business analytics workloads have evolved over time to require significant processing power, memory and I/O bandwidth. The mainframe is designed to provide a balanced system optimized for sophisticated new workloads, such as data mining and text analytics of both structured and unstructured data. The cache size and structure, the internal bandwidth, and the I/O structure and bandwidth are more robust when compared to CPU speeds on other platforms.
- Parallel processing – the mainframe environment for BA leverages popular capabilities, such as parallel processing, without requiring any design changes and simultaneously supports batch and real-time operations. The system administrator can assign one or more system processors for the exclusive use of an LPAR, or allow all processors to be used on some or all LPARs. The mainframe's zIIP and zAAP engines are also well suited to handle parallel queries (with

an emphasis on star schema designs) common in BA workloads at a competitive price/performance point.

- Smart query management – having a single, centralized copy of data accessed by all users makes sense, but can be a nightmare if you cannot control and differentiate between the critical work and less important queries in the system. IBM Smart Analytics Optimizer, available on System z, addresses the challenge of managing and tuning databases for long-running queries that do not allow for ad hoc processing. IBM Smart Analytics Optimizer accelerates queries through in-memory processing, massively parallel architecture, row and columnar store technologies, highly compressed data, and compressed data operations. Smart Analytics Optimizer is breakthrough technology, developed at the IBM Labs to dramatically accelerate queries. It is a functional extension to DB2 where DB2 performs tasks such as query prioritization and SQL parsing. The Smart Analytics Optimizer processes queries or portions of queries that the DB2 optimizer deems to be more effectively run on the Smart Analytics Optimizer rather than execute it itself. There are no application changes or tuning required to the SQL. Companies are now able to run queries 100 to 1,000 times faster than they have been able to do in the past.
- Lower total cost of ownership than on distributed systems – because it is run from a centrally managed platform, System z helps alleviate the expenses that would go with supporting business analytics workloads across multiple servers, including redundant copies of software, server maintenance, systems administration, monitoring, and backup. The power consumption seen on multiple servers can also be significantly reduced, reducing energy costs and providing for a “greener” operation.
- Open standards – System z supports all leading open standards and languages that enable it to integrate with all of today's leading BA solutions and platforms. The platform runs Linux and Unix-based applications, and supports interfaces to other applications via service-oriented architecture. No matter what skill set may be resident in the enterprise, the System z can support it.
- Ironclad security – the System z platform secures data with cryptography, encryption, user identification and

authentication at all levels. The mainframe's z/OS operating system has the highest security rating from the government. Plus, it is easier and more cost effective to secure BA data in a centralized location such as System z than across distributed environments.

Deployment Options

Deployments of BA environments have changed radically in the past five years. Where once, everything had to be created in-house within a company's firewalls, we now see more flexible deployment options. This can be good news and bad news: the good news is that companies today can use many different BA products to quickly create analytic applications. It gives enterprises the flexibility to start where you need from a functionality perspective: simple reporting or querying, multi-dimensional comparisons, or advanced analytics like predictive or statistical analytics. However, the bad news is that these can lead to disjointed, non-integrated sets of data and applications (silos of analytic solutions). The rationales to create these potentially chaotic environments include perceived lower costs of these environments, rapid time to implement, and, of course, demands from powerful factions within the organization for their "own" BA solutions. Yes, we need to support all our information workers BUT not at the expense of a maintainable and sustainable BA architecture.

Unfortunately these disjointed "environments" actually cost more in aggregate. They also fracture the data so badly that consistency and reliability are diminished if not destroyed completely, and make maintenance far more difficult. Many organizations are in the throes of consolidating and standardizing their disjointed environments to reduce overall costs and extend BA to an even larger audience, expanding it to encompass consumers and collaborators, not just producers. As they go through this process, they find that BA on the System z is a natural choice for consolidation since significant percentages of data already reside on the mainframe and many information workers are already familiar with the technology through the operational activities they already perform there. Why not eliminate the problem of training and retraining information workers on different interfaces, and disconnecting BA from the operational workflows? The System z supports the co-location of BA and operational applications easily, as well as reducing the costs of integrating non-mainframe data into the BA environment, thus making decision-making easily embedded into mainframe

operational activities.

Once the deployment options have been chosen, the mainframe supports an infrastructure for today as well as for the future:

- Private cloud computing for BA – For example, IBM's Smart Analytics Cloud enables the delivery of BI and analytics at a customer's location through a private cloud deployment. This leverages the hardware, software and services offered to support a large multi-tenant or multi-line of business base of information workers. Customers on the system z benefit from the massive scale, superior virtualization capabilities, and reliability, availability and security that are hallmarks of that environment. In addition, IBM Services offers a range of BA services designed to create this architecture and to enable knowledge sharing across the diverse worker base.
- Turnkey solutions where the data warehouse and BA applications are shipped – ready-to-go – in a single box from the vendor. The largest expense item for many BA environments is the actual labor it takes to implement these solutions. By purchasing a turnkey solution like the IBM Smart Analytics System, the labor costs, time, and effort are reduced to a minimal amount. Enterprises considering the System z can now get the entire suite of data warehouse enabling and optimization tools plus a robust BA infrastructure co-resident with the data warehouse and operational applications.
- Do it yourself construction of BA applications – even for those organizations that want to do it all themselves, having a single environment with common interfaces and shared services / procedures is a great productivity boost. See the next section for more technological considerations for do-it-yourselfers.

Summary

Until recently, business analytics solutions were only used by a relatively small handful of analytically skilled professionals in organizations. However, in today's highly competitive business landscape, BA needs to be part of everyone's job. Whether producing or consuming information, every employee is an information worker, and needs appropriate access to the right information to make the best decisions.

In addition, an emerging class of information worker – information collaborator – demands new, more agile and responsive BA technologies. Information collaborators understand the power of real-time or near real-time data, and employ structured and unstructured data, collaboration technology, and social media to further extend knowledge throughout the enterprise. Cognos 10 includes many new features that embrace these requirements, allowing BA implementers to expand BA by including this new group of information workers.

This calls for an information technology infrastructure that provides a powerful and agile environment that weaves actionable information into the day-to-day fabric of end-users' decision making. With System z, organizations have access to a well-managed, cost-effective, scalable, and responsive BA environment that supports all workloads and all requirements of today's diverse knowledge workers. The mainframe is ahead of distributed platforms in being functionally rich, with sophisticated features that ensure the highest application availability and the timeliest information. System z is the right solution to meet the fast-changing needs of today's more enlightened and demanding information worker.