Business Intelligence at University of Dayton

Anthony DeCerbo
Meaghan Duffy
Steve Smith
Warren Scoville

MBA 664
Winter 2009
History of Business Intelligence

Business Intelligence allows people at all levels of an organization to access, interact with, and analyze data to manage the business, improve performances, discover opportunities, and operate efficiently. Over the past two centuries Business Intelligence has evolved to aid businesses in handling their most important data in the most efficient ways possible. Business Intelligence has evolved from major milestones of the forerunners of BI solutions (Howson 2008). Many tools of the past have helped to form and structure Business intelligence and why it is the way it is today. Rapid change of technology has been one driver of the hectic pace of business change it also has allowed business intelligence for everyone, not just information technology experts, programmers, and power users. The technology and BI tools have improved over time to extend the reach of business intelligence (Biere, 2003).

The 1970s was considered the Early-End User Computing era. During this time, user-friendly languages allowed end users and the unfriendly IT environment to understand each other and work together in a more appropriate manner for both parties. At this stage in time, end users had to wait for systems, wait for programming changes, and wait for reports. During these times, the end user might have waited a month for complex, printed report that ran on a mainframe computer. There was a lot of time being wasted while the end users waiting for reports to emerge from central programming and computer sites. The early tools that were used were all “do-it-yourself” solutions. At the time, these solutions seemed like a good idea, but with Vendors setting up tools and formats to do so lead to extremely complicating matters (Biere, 2003).

During the 1980s the Information Center concept had formed. The information center was a central support organization designed to provide services for end user and to act as relationship between non-technical user and IT. The Information center supplied end users with the necessary tools to help with the required data for their organization. The main reason for these centers were to provide means for end users to increase their productivity instead of having to wait and simply do nothing until they received their data to continue (Biere, 2003). Also the relational databases were being used during the 1980s. This was when data was stored in tables with rows and columns, completely different from what the excel spreadsheets look like today. The relational databases became more convenient for end users to use, but still was complex with the amount of data they needed to strain together to obtain the needed information to create the
report. At this time, reports could be run by end users as long as they were simple, but still needed to be run at off times to not interfere with others work by tying up the system.

Throughout the late 1980s business migrated from the mainframe computers to client servers. Some of the main reasons for leaning toward client servers were due to the expensive cost of mainframe, the idea that data should be stored on smaller, compact, and less expensive boxes, and logic and calculations should take place under the control of the end user. During this era of client server systems, organizations realized that reengineering data in Business Intelligent forms and formats were more ideal to the company and the most common form was a relational store that supported SQL (Biere 2003).

During this time, personal computers were becoming popular within the organizations. Personal Computer opened many opportunities for employees and they were able to work on their own without the aid of IT. The office applications became popular and the personal computers allowed end user to develop their own application and present the data in their own form that was unique and meaningful to them and their organization. The Office applications, which included word, excel, and access, allowed the users to create resourceful reports that they were never able to do before (Biere, 2003). Charts, graphs, and grids were created to show their collection of data and their reports in the best way possible to be understood specifically by the company in need of the information.

Throughout the 1990s the need for Data warehousing became an important tool for Business Intelligence. Data warehouse is a collection of data removed from a variety of operational systems, transformed to make the data consistent, and loaded for analysis (Howson, 2008). Data warehousing was an accurate way to store data in an organized way. It offered the organization creative ways to present its data and to analyze the data toward a specific request. It allows the company to conduct a more detailed analysis of the topic that the company is trying to focus on.

In the early 1990s the Internet became exceptionally popular. The web-based Business Intelligence has allowed tools to be organized across corporate intranets and extranets to thousands of employees and external customers in a matter of hours not months. With this client/server computing of the early 1990s, it took days to install and configure PCs for just a handful of users. The web has simultaneously broadened the reach of Business Intelligence (Howson, 2008).
The road to implementing a more accurate and consistent Business Intelligence into an organization was long and tedious process. Over the years, the collection of data has traveled through different eras and has improved in many ways to become more convenient for the organization. It all started with the early end user era where organizations used do-it-yourself solutions to process data, information centers was the next step where central support was offered to the organizations, which reduced time waiting and increasing productivity. Soon later mainframes were transformed into client server systems that were easier to deal with, and the idea of data warehousing was implemented shortly after this. All of these eras have contributed to improving Business Intelligence to accurately handle the most important data of the organization today.

**Description of Business Intelligence**

Business intelligence has become essential at the corporate level. Kelly (2005) found that spreadsheet-based processes cannot address centralized data, integrated performance metrics or timely controls. A business warehouse allows standardized reports for casual users while experienced end users to customize reports as their requirements change. And while it is important for companies to understand what has happened in the past, it is very important to do strategic modeling to gain a better understanding of how to move forward. Business warehouses can vary in scope from a major corporation’s ERP supported central data warehouse costing millions of dollars to implement down to a small firm’s mini-data mart that came on line for less than $25,000 (Miller, 2008). Further data warehousing can be obtained by small firms through the use of optional add-on toolsets. IBM has done this for its online transaction processing server (Kanaracus, 2009b).

Scorecards or corporate performance management is the process, methodology, technologies and metrics that enable an organization to manage its business performance (Kelly, 2005). It replaces the excel-based spreadsheet form of management that can days to generate and is open to data entry errors. Scorecards integrate data from different sources including spreadsheets, databases and flat files. They can be used to filter through the high volume of data and information to allow management to understand the critical measures driving the organization’s operations from the strategic corporate-wide context down to the managers who
influence daily business. The focus should be on the visionary metrics at the top. Subsequent metrics should assist with achieving those top metrics (Pauli, 2009).

A Data Cube is an array of values used to describe a time series of data in a database. It can vary from a two dimensional spreadsheet of values taken over time to a multi-dimensional data hypercube. The user can specify what dimensions or subsets of data they want to look at and change the data’s orientation. Common operations include: drill down from summarized data to more detailed data, pivot the dimensions of a report, and slice and dice a data cube by one or more dimensions. This can allow the user to identify hidden trends and clusters, and to project future actions or results based on historical information. The user can model the data by changing a dimension and examining the results.

Dashboards are user interface tools designed to allow the non-technical end user the ability to make standard queries with minimal effort. End users are more likely to embrace business intelligence tools that they can use independently (Daniel, 2008b). Executive dashboards are the most frequently used, although the illustration below is for an unclassified engineering group dashboard. Some of the dropdowns have been activated to show the variety of information that can be accessed.

![ARI Dashboard](ARI_Data_Warehouse.png)

**Where Business Intelligence is going**
Organizations need a new generation of knowledge-enabled systems that can give them a tool to clean, store organize and dispense not only the source data and information, but the value-added knowledge of the organization (Chen, et.al. 2009). They found that workers will need to freely extract the needed data, analyze it in the context of different decision support modules, use changing parameter sets, and broadcast the value-added knowledge to the clients. The current economy is forcing firms to pursue profit by digging deeper into the information already at their fingertips (King, 2009). Some software can bring in external data to better understand how economic trends are affecting the customer and the company.

Corporate Performance Management enables businesses to satisfy today’s compliance conscious environment. It uses tools to drive operational performance, enterprise scorecards to monitor performance to the plan, and enterprise business intelligence for reporting and analysis. The package is designed to allow compliance to Sarbanes-Oxley and Basel II (Schauer, 2004). As the world becomes hotter, flatter, and more crowded companies need an ever greater breadth of information to remain competitive. Successful companies need to have access to highly relevant information so that decisions can be made in real time (Caputo, 2009).

Social media is the collection of “online applications, platforms and media which aim to facilitate interaction, collaboration and the sharing of content” (McCann, 2008). This content can be chat rooms, blogs, twitters, shared photos videos or music, and social networking. This has lead to a greater willingness to share information as opposed to the information silos found in the past. One of the biggest examples of social media is the on-line dating services. They combine large databases (up to 4 terabytes) with psychological profiling, business intelligence, algorithms, and communications technologies (Mitchel, 2009).

Shared information can facilitate connections throughout an organization and can lead to better decisions, productivity increases, reduction in product introduction time, and cost savings. One of our firms encourages IM-ing, has frequent on-line interviews with executives, open Q&A streaming, open documentation on SharePoint sites and ERP supported data warehousing for engineering, finance, production control, logistics, marketing, and sales. Unfortunately learning to share can be difficult for organizations and managers that are accustomed to a high degree of autonomy (Kelly, 2005). But the ability to enhance historical data with relevant and timely information from the Web and shared information is becoming critical for thorough analysis of information (Daniel, 2008b).
Ontology is the set of representations used to model a domain of knowledge. They can include attributes and relationships with information about their meaning relevant for modeling a domain. Ontologies are used for integrating heterogeneous databases (Gruber, 2008). The metadata taxonomy is used to map the data into the business warehouse. Having the users determine their tags is one method of providing a multi-level structure to the database.

The Semantic Web is the next step in information searches. It is designed to provide the context of the information in addition to the information. It will provide a ‘common, standardized framework that allows data to be shared and reused across applications, enterprises and community boundaries” (Caputo, 2009). The search query will be able to understand the connections between apparently unrelated pieces of information and draw conclusions relevant to the end user. Ontologies are part of the standard stack required for the Semantic Web. They use a standard conceptual vocabulary for data modeling so that data can be translated, exported as required across independent systems (Gruber, 2008). The five layer model consists of the following:

- Resource layer – data can be structured or unstructured, and can be dynamic or static.
- Data conversion layer – enterprise metadata management ontology and the extraction, transformation and loading module are used to map the various databases into the data warehouse.
- Data storage and data management layer – create and manage data cubes, user accounts, and calendar management.
- Knowledge/trend/pattern layer – manipulate data cubes, data mining, and statistical modules.
- User process layer – reports, web functions, etc. (Cheng, et. al., 2009)

The separation of strategic and operational decision making is eroding as it (BI) gets integrated with workflow, business automation, and proactive systems (Humphrey, McKnight, 2004). Higher management is pushing the information flow down into middle management to enhance the tactical decision making process which can trigger real-time collaboration and decision making. Supply chain management can do cost benefit modeling to determine the transportation modality required from delinquent suppliers to support the manufacturing floor’s sequencing requirements (Lee, et. al. 2009). “Operational business intelligence can provide:
• A flexible framework for satisfying informational needs…
• Detailed timely information to the users
• Faster query response time
• More ad hoc querying capability
• Access to lowest granularity data
• More user friendly formats for quick operational decisions
• A scalable model that can roll up to the executive dashboards or scorecards” (Joshi, 2007)

Care must be taken to ensure that the business intelligence system’s performance does not get bogged down by the operational business intelligence’s requirements for high volume timely data. One of the most promising areas for business intelligence is the new tools that automate data collection, assembly and delivery processes (Daniel, 2008b).

Relevance of Business Intelligence to institutions of higher learning

Typically when the subject of business intelligence is brought up, it is in a corporate context. A majority of business intelligence implementations have been in a corporate setting, where reports are ran about year-to-date sales, and decisions are made based on sales forecasts. Business intelligence in a higher education setting does share some aspects of BI in a corporate environment, but there are other aspects that are unique to an education environment. There are reports and decisions made around the university’s income, and expenses that affect the bottom-line, but there are many additional uses of BI in higher education. Business intelligence reports can help with overall “institutional intelligence”. Higher education does lag behind corporations in their adoption of business intelligence with only 29% of institutions reporting that they currently used BI in a May 2008 survey. (McClure, 2008) With an ever increasing amount of data for administrators to look through, and the need to make better decisions more and more school are looking to implement BI. According to same survey done in May 2008, 51% of all institutions surveyed were looking to implement a BI solution within 6 to 24 months.

Some of the easiest reports to think of that can help university administrators make decisions are based on data from departments such as the admissions office and university advancement. Data about current student enrollment such as how many people are in a particular major and grade level can help plan how many classes and sections should be offered, what class
sizes should be, and help schedule rooms. Currently a lot of these decisions are based off data and reports from a previous semester or year, and are made a semester or so in advance. One potential use of BI in this area would be to look at trends from data from years past and predict a need for additional professors, or classrooms for a particular class based on an increasing interest in a subject area. This type of use of BI would allow administrators time to plan class meeting times and rooms, or to even find additional professors, and avoid a last minute scramble to find these additional resources. BI reports and data could also be tied with new student recruitment campaigns. This data would allow the admissions office to see which campaigns are working, and which are not. Another example we have talked about in class before, would be the ability for the admissions office to track interests and deposits for a particular major, and align their campaigns to try and bring in a class that balances across all of the university’s departments instead of overloading or neglecting a particular department.

While the use of business intelligence would help administrators make informed decisions about admissions and funding campaigns, it would also help answer questions such as are we serving the students the best we can. It is one thing to use BI to predict how many students will be taking a particular class in a couple of semesters to make sure all of the needed resources are in place, and this is one way of serving the students by making sure there is enough space for everyone who will need the class. Looking at that report, and making decisions about the number of sections, and how many seats are available is only one step though. By looking at trends with BI, administrators can look and see if a particular class is being offered at a good time. A report might show that students who have taken Biology 320 and Chemistry 240 during the first semester and Biology 410 and Chemistry 300 during the second semester tend to graduate in 4 years, where if only one of those classes is taken in the second semester they tend to graduate in 5 years. If both of those classes are only offered in once during the second semester, and occur at the same time that might not be in the best interests of the students. The registrar’s office could then take that data, and move classes around so that students are able to take both classes the second semester. Many schools have also started to use their BI solutions to identify and report on key indicators for student success. Different office around the universities then use these reports to try and identify students that may be at risk of failing, and offer them additional help and resources.
Institutions of higher education are lagging behind many corporations in implementing data warehouses and business intelligence solutions, but more and more universities are starting to investigate BI. There are obvious benefits to the business side of the university by reporting on enrollment numbers and financial numbers, which will help the university, perform at its best while attempting to keep costs as low as possible. There are also many other benefits to a BI solution, including making sure that a university is doing its best to serve the students the best way possible.

**Short-comings of UD’s current BI and data warehousing solution**

The University of Dayton is currently using a Teradata database for a data warehouse, and Cognos ReportNet 1.1 for a BI reporting solution. The data warehouse stores information from Colleague, our student information system. The Teradata warehouse was put into place a number of years ago, to keep track of data that is purged from Colleague on a semester basis. For example, if a student drops a class that information is stored in our student system for that semester, but is purged at the end of the semester. ReportNet is used to make reports from any of the data that is stored in the data warehouse. Both systems are used by several different individuals from different departments throughout the University for reporting.

Both Teradata and ReportNet do suffer from a couple of short-comings though. When the data warehouse was put into place, it was designed specifically to receive information from the student information system, as was the extract, transform, and load process that move data between the two systems. Right now there are several different information systems on campus, but the only data that is put into the warehouse is student information, which means that currently no one can use the university’s BI solution to report on other business, financial, or admissions information. Also, even though we call the Teradata database a data warehouse it is more like an operational data store. The data is used more for daily reporting, than it is for trend analysis.

Currently, neither Teradata nor ReportNet are being used for data mining or trend analysis. The majority of the reports that are being ran on the current system offer information about faculty workloads, and class size. For example, there are reports that will output a class roster but there are not reports to determine if that class is being offered at a good time. When the Teradata and ReportNet systems were put into place, they were both IT department projects
and did not have an executive sponsor. When someone calls IT wanting certain information, we steer them in direction of ReportNet and will often write a report for them. The lack of a sponsor means that there is no one forcing ReportNet as the BI solution across the entire university. There are also very few people that are “power users” writing their own reports, or using the ad-hoc reporting capabilities of ReportNet, which means that people must ask the IT department to write these reports for them.

While some of the short-comings of these systems are technical, there are also a couple that are organizational. Many of the organizational problems are just getting people inside of the university to change the way they currently do business, and adapt to a new way. The IT department can make this easier by asking for input from people outside of the IT department about what data and reports are needed, as well as by offering training about how to use these reporting solutions so that people can write their own reports and queries. The data warehouse and ETL processes could be redesigned to pull more information in, and offer additional data to report on and draw conclusions from. Correcting both problems though can be quite challenging. The University of Dayton is currently in the middle of implementing and new enterprise resource planning system, which will hopefully help fix a lot of these problems.

**Benefits of new ERP based data warehouse and BI solution**

Successful implementations of large-scale business intelligence projects follow the same principles found in ERP projects. There must be an executive champion to drive the project. Subject Matter Power Users must be included in the current-state future-state gap analysis of processes, data cleansing, data transformations and removal. Technology expert are required to determine the technological gap analysis. These gaps have to be prioritized and determination made at the executive level on what will be in-scope. The master data has to be organized around key aspects of the organization and be captured, stored and processed in a consistent manner (Kanaraacus, 2009a). The implementation is controlled through a project blueprint that covers the footprint and functional depth of all in-scope functions and is managed through milestone gated reviews. Areas that can make or break an implementation include: data cleansing, metadata management, change management, and end user training (Pauli, 2009).

The university is currently implementing a new ERP system that will combine all of our human resources, financial, student, and admissions information into one system which will
eliminate many of the disparate systems throughout the different departments in the university. The university is also going to implement a new operational data store, enterprise data warehouse, and reporting solution during the ERP project. The ODS and EDW are from the same company that writes the ERP software, and all of the data from the operational system will be moved into the ODS and EDW systems. So people throughout the university will be able to report on more than just student data. In addition to implementing a new ODS and EDW, part of the new ERP solution will also include upgrading our reporting software to Cognos 8 which includes more features than the previous version the university is currently running.

The information pyramid

- Information needs to reach all levels of campus.
- Data from lower levels must be transformed to upper levels.

(SunGard Higher Education, 2009)

Since the new ODS, EDW, and reporting software will be part of the overall new ERP plan, there is now several executive sponsors for the project. There has been a lot of communicate between the IT department, and various other departments throughout campus about how the new systems should be setup, what data needs to be kept, and what types of reports should be written. Also, people realize that they way they do certain things currently will need to change once the new ERP systems is completely in place and ready to use. With some of the problems the previous data warehouse and BI report server had, and the amount of work it
would have taken to fix those problems, tying it into an upgrade such as the installation of a new ERP was the best time to fix those shortcomings. The new operational data store, enterprise data warehouse, and business intelligence solution will be designed to deliver the right information to the correct jobs that need it.

**How the new data warehouse is setup**

The University of Dayton’s new ERP system, Banner, will be setup with three different database back-ends, an operational system, an operational data store, and an enterprise data warehouse. The operational system, Banner, will be where the majority of daily activities will take place, such as entering grades, entering purchase orders, and updating employee records. All of the tables in Banner have been designed to normalize all of the data, remove duplicate data, and optimize the system for reading and writing to the database. While this design is great as an operational system, it is not the best for reporting. The second database that will be part of the new ERP implementation will be the ODS.

**Banner ODS Architecture and Design**

(SunGard Higher Education, 2009)

Banner’s operational data store has been designed specifically for reporting, and is installed on a different server than the operational system. Since the database is installed on a separate server, it lessens the load on the operational system. The ODS uses Oracle Warehouse Builder for the extract, transform, and load process which runs on a scheduled basis. When data
is pulled from Banner to put into the ODS, it is denormalized and put into composite table to optimize it for reporting purposes. Denormalizing the data, and putting it into composite tables speeds up the time it takes to run reports by eliminating many smaller tables and putting the data into one large table which eliminates the need for a lot of expensive table joins.

After data is moved to the ODS, it is moved to a third database called the enterprise data warehouse by the ETL process on a scheduled basis. The EDW groups and summarizes the data from the ODS into different star-schemas. The EDW is optimized for reporting, just like the ODS, but is specifically used to draw summaries and trends from. Some of the details that are stored in the operational system and operational data stores are not kept in the data warehouse. For example, in Banner and the ODS, someone could pull a roster of all 35 students that are enrolled in Biology 101. However, in the EDW, a user would only see summary data about Biology 101 such as 40 students originally signed up for the class, but only 35 students completed the course.
References


