

Why You Need a Data Warehouse

A Special Report for mid-sized organizations

Revised and Updated

Joseph Guerra, Vice President & Chief Architect

David Andrews, President



700 West Johnson Avenue
Cheshire, CT 06410
800.775.4261

www.andrewscg.com

www.rapiddecision.net

www.jdedwardsadvisor.com

Chances are that you have heard of data warehousing but are a little fuzzy on exactly how it works and whether your organization needs it. It is also highly likely that once you fully understand exactly what a data warehouse can do, you will decide that one is needed.

Data warehouses are widely used within the largest and most complex businesses in the world. Use within moderately large organizations, even those with more than 1,000 employees remains surprisingly low at the moment. We are confident that use of this technology will grow dramatically in the next few years, even taking the current economic downturn into consideration.

In challenging times good decision-making becomes critical. The best decisions are made when all the relevant data available is taken into consideration. The best possible source for that data is a well-designed data warehouse.

The concept of data warehousing is deceptively simple. Data is extracted periodically from the applications that support business processes and copied onto special dedicated computers. There it is validated, reformatted, reorganized, summarized, restructured, and supplemented with data from other sources. The resulting data warehouse becomes the main source of information for report generation, analysis, and presentation through ad hoc reports, portals, and dashboards.

Building data warehouses used to be difficult. Many early adopters found it to be costly, time consuming, and resource intensive. Over the years, it has earned a reputation for being risky. This is especially true for those who have tried to build data warehouses themselves without the help of real experts.

Fortunately, it is usually no longer necessary to custom build your own data warehouse. The heavy

lifting has already been done by others. Prebuilt solutions are now available that dramatically reduce the effort and risk. As a result, the time has come for mid-sized organizations to develop a thorough understanding of data warehousing. The goal of this paper is to take the mystery out of this subject and present the case for data warehousing in simple business terms. What you will learn is likely to pleasantly surprise you.

The Business Case for Business Intelligence

Managing a business has never been easy. Trying economic conditions make it even harder. The challenge is to do more with less and to make better decisions than competitors. Having immediate access to current, actionable information can make a huge difference. It is the reason why spending on Business Intelligence (BI) solutions has continued to grow during a severe economic downturn.

Executives cannot afford to make decisions based on financial statements that compare last month's results to a budget created up to a year ago. They need information that helps them quickly answer the basic questions: What still works? What continues to sell? How can cash be conserved? What costs can be cut without causing long-term harm?

Business intelligence systems provide the ability to answer these critical questions by turning the massive amount of data from operational systems into a format that is current, actionable, and easy to understand. They provide information that is decision ready. They allow you to analyze current and long-term trends, alert you instantly to opportunities and problems, and give you continuous feedback on the effectiveness of your decisions.

BI is not a new idea. The largest and best-managed organizations in the world have been making use of it for more than a decade. Along the way, the giants that pioneered BI made an important discovery – *the path to true business intelligence passes through a data warehouse.*

The Connection between Data Warehousing and Business Intelligence

The Data Warehousing Institute defines business intelligence as:

The process, technologies, and tools needed to turn data into information, information into knowledge, and knowledge into plans that drive profitable business action. Business intelligence encompasses data warehousing, business analytic tools, and content/ knowledge management.

The fact that the leading authority on BI calls itself the Data Warehousing Institute highlights the vital role that data warehouses play. Unfortunately, the warehouse adds its value behind the scenes. Its job is to provide data to the high-profile tools and applications with which users interact. This background role can hide its significance, especially since BI solution vendors frequently play down the importance of the data warehouse.

Technically, it is not necessary to build a data warehouse in order to create a BI environment. As a result, there are many substandard solutions on the market that avoid the use of data warehouses. Those advocating these solutions often suggest that the absence of a data warehouse is a good thing. They

are following the old marketing adage, “If you can’t fix it, feature it.” What countless BI pioneers have discovered, however, is that taking the short cut around data warehousing will put you on a path that leads to lost time and money.

Data Warehouse Basics

The concept of data warehousing is not hard to understand. The notion is to create a permanent storage space for the data needed to support reporting, analysis, and other BI functions. On the surface, it may seem wasteful to store data in more than one place. The advantages, however more than justify the effort and cost of doing so.

Data warehouses typically:

- Reside on computers dedicated to this function
- Run on a database management system (DBMS) such as those from Oracle, Microsoft, or IBM
- Retain data for long periods of time
- Consolidate data obtained from many sources
- Are built around a carefully designed data model that transforms production data from a high speed data entry design to one that supports high speed retrieval.

An extract, transform, and load (ETL) software tool is used to obtain data from each appropriate source, including whatever ERP systems are in use. ETL tools read data from each source application, edit it, assign easy-to-understand names to each field, and then organize the data in a way that facilitates analysis.

A key aspect of the reorganization process is the creation of areas within the warehouse that are referred to as data marts. In most cases, individual data marts contain data from a single subject area

such as the general ledger or perhaps the sales function.

The best data warehouses do some predigesting of the raw data in anticipation of the types of reports and inquiries that will be requested. This is done by developing and storing metadata (i.e., new fields such as averages, summaries, and deviations that are derived from the source data). There is some art involved in knowing what kinds of metadata will be useful in support of reporting and analysis. The best data warehouses include a rich variety of useful metadata fields.

The most difficult thing about creating a good data warehouse is the design of the data model around which it will be built. Decisions need to be made as to the names to give each field (since the names in application databases can be obtuse), whether each data field needs to be reformatted, and what metadata fields should be calculated and added. It is also necessary to decide what, if any, data items from sources outside of the application database should be added to the data model.

Once a data warehouse is made operational, it is important that the data model remain stable. If it does not, then reports created from that data will need to be changed whenever the data model changes. New data fields and metadata need to be added over time in a way that does not require reports to be rewritten.

The value a Data Warehouse based BI solution provides

Once a data warehouse is in place and populated with valuable data, good things begin to happen. Examples of the many ways in which data warehouse-based BI systems deliver value to their users include:

- *The generation of scheduled reports.* Moving the creation of reports to a BI system increases consistency and accuracy and often reduces cost. A greater number of more useful reports result from the power and capability of BI tools. The creation of reports directly by end users is easier to accomplish in a BI environment.
- *Packaged analytical applications* A growing number of outstanding analytical software applications are coming onto the market. These packages provide predefined reports and metrics that business units can use to measure their performance.
- *Ad hoc reporting and analysis.* Since the data warehouse eliminates the need for BI tools to compete with transaction processing, users can analyze data faster and generate reports more easily. The tools that come with BI systems also tend to vastly improve the analysis function.
- *Dynamic presentation through dashboards.* A growing number of managers want access to an interactive display of up-to-date critical management data. Sophisticated displays that show real time information in creative, highly graphical form are often called dashboards. The name comes from the similarity to the instrument panel on an automobile.
- *Drill down capability.* The leading BI systems all allow users to drill down into the details underlying the summaries in reports and dashboards. The presence of a data warehouse makes it practical to use this capability as much as needed. Without one, limits need to be placed on user access to application data.
- *Support for regulations.* Sarbanes-Oxley and related regulations create demands that transaction systems cannot always support. Well-designed BI solutions can ensure that the necessary data is retained in the data warehouse for as long as is required by law.
- *The creation of metadata.* Data warehouses sit between source applications and BI tools, creating an ideal opportunity to predigest some of the data. Metadata is defined as “data about data.” It can include something as simple as an average. Data warehouses can be used to create and store a great deal of metadata of potentially great value.
- *Support for operational processes.* The creation of a sound BI infrastructure is often the best way to meet certain ongoing business needs. The most common example is to facilitate the consolidation of financial results within complex organizations, especially those whose divisions use different software systems. Meeting regulatory reporting requirements is another common situation.
- *Data mining.* The outstanding software tools that can sift through mountains of data and uncover hidden insights work best off a data warehouse.
- *Security.* A data warehouse makes it much easier to provide secure access to those that have a legitimate need to specific data and to exclude others.

This long list of benefits is what makes BI based on a data warehousing an essential management tool for businesses that have reached a certain level of complexity.

Business Intelligence Analytic Tools

Analytical tools are the rock stars of the BI universe because they provide the dashboards, drill down capabilities, and trend analysis needed by users. Most people get their first introduction to BI by seeing a demonstration of one of these tools. Such demonstrations can be compelling, especially when shown to end users and executives who yearn for better reporting, analysis, and inquiry capabilities.

Outstanding products are available from many different vendors. Market leaders include Business Objects (an SAP company), Cognos (an IBM company), Oracle, Microsoft, and MicroStrategy.

An evaluation of the pros and cons of the various offerings is beyond the scope of this paper. All of them are good enough to form the foundation for a successful BI solution.

The vendors who sell these analytical tools naturally focus attention on the benefits that result from their use. They do not always make it clear that a prerequisite for their effective use is a properly designed data warehouse. When asked directly, the leading vendors will all admit that their tools work best when extracting data from a warehouse.

Is a Data Warehouse really necessary?

The data needed to provide reports, dashboards, analytic applications and ad hoc queries all exists within the set of production applications that support your organization. Why not just use one of the BI tools to obtain it directly?

Countless BI pioneers have discovered the hard way that the “direct access” approach simply does not work very well. There is no reason to add the name of your organization to the long list of failures.

Some of the many reasons why direct connection almost never works include:

- New releases of application software frequently introduce changes that make it necessary to rewrite and test reports.
- These changes make it difficult to create and maintain reports that summarize any data that originates within more than one release.
- Field names are often hard to decipher. Some are just meaningless strings of characters.
- Application data is often stored in odd formats such as Century Julian dates and numbers without decimal points.
- Tables can be structured to optimize data entry and validation performance, making them hard to use for retrieval and analysis.
- There is no good way to incorporate worthwhile data from other sources into the database of a particular application.
- Developing and storing metadata is an awkward process without a data warehouse – there is no obvious place to put it.
- Many data fields that users are accustomed to seeing on display screens are not present within the database, such as rolled-up general ledger balances.
- Priority always needs to be given to transaction processing. Reporting and analysis functions tend to perform poorly when run on the hardware that handles transactions.
- There is a risk that BI users might misuse or corrupt the transaction data.
- There are many ways in which BI users can inadvertently slow the performance of applications.

Even though additional hardware and software are needed, the presence of a data warehouse costs less and delivers more value than direct connection. Every year predictable declines in the cost of computer processing and storage make the case even stronger.

The declining cost of hardware not only makes BI solutions cost less, but also encourages businesses to retain even more data on which BI solutions can feast. Every year data warehousing becomes even more economically attractive.

Data Warehouse: build versus buy?

Three options are available to obtain a data warehouse:

- Buy it completely prebuilt
- Build it using frameworks
- Create a custom warehouse

The easiest and most cost effective way to get the data warehouse you need is to buy it. Numerous options are available, especially for the most popular ERP suites including those from SAP and Oracle.

The better pre-built solutions make it possible to have a data warehouse up and running quickly, sometimes in just a few weeks. It is often possible to see a live demonstration using your own data before committing to its purchase.

When a complete pre-built data warehouse solution is not available it is sometimes possible to buy frameworks from which a data warehouse can be created. This approach involves less time, cost, and risk than building your own warehouse.

In theory, frameworks can be customized to your special needs. However, doing so is not always easy, inexpensive, or foolproof. Success is highly

dependent on the skills of those doing the customization. Ongoing maintenance and support can be a major headache for those that choose to use frameworks.

The least attractive option is to create a completely custom data warehouse. Doing so can sound simple – obtain an ETL tool and use it to move data from application databases into a data warehouse. In practice, this option turns out to be far more complicated. A significant percentage of those that have tried to do it themselves have either failed or spent far more money and time than planned.

Many companies have spent more than a million dollars over the course of a year or more without being satisfied with the result. High ongoing maintenance costs never go away in this case.

Those exploring the options for obtaining a data warehouse need to be aware that custom creation is sometimes strongly advocated by vendors that hope to generate additional income by doing so. The available prebuilt options should always be examined before trying to build it yourself or contracting to have it done for you. The bottom line is that friends don't let friends try to build their own data warehouses when there is no need to do so.

Evaluating pre-built Data Warehouses

A growing number of fully and partly built data warehouses have become available. When more than one option is available it is necessary to evaluate the pros and cons of each. Obviously the first concern is whether the applications you use are fully supported including every version, module and release in use.

Other criteria to consider include:

- *Time to value.* How long does it take to get to get the warehouse operational?
- *Design and stability of the data model.* Nothing is more important than this.
- *Quality of metadata.* Well-designed metadata that is enhanced over time can be invaluable.
- *Availability of analytics and reports.* The best options include useful prebuilt reports, dashboards and analytics.
- *Real time updating.* How often is the data warehouse synchronized?
- *ERP performance.* Is the impact of data movement on ERP performance minimized?
- *Which BI tools are supported?* Is one you already own or prefer available?
- *Track record.* Have others enjoyed success with the data warehouse being offered?
- *Service offerings.* What kind of help is provided and how good is the reputation of those providing it?
- *Quality of maintenance and ongoing support.* Will you be on your own after initial installation? Do improvements come on a regular basis?
- *Technical environment.* Which servers, operating systems, databases and other tools are supported?
- *Performance.* Design choices can have a huge impact on performance and hardware cost.

Why Start Now?

The value of a data warehouse increases over time. It therefore pays to start putting one in place as soon as practical. Those who delay starting down this path could remain at a disadvantage versus competitors that begin sooner. Payback in terms of business benefits comes rapidly in many ways:

- Hard savings come from things like discovering lost discounts in payables or that sales people are offering discounts beyond approved limits.
- Real-time consolidation of financial data becomes practical.
- Debates cease over which source of data is correct.
- The IT costs and staff dedicated to reporting are greatly reduced.
- Use of BI becomes more widespread as users discover its value.
- Users become more skilled at extracting useful information with experience.
- Historical data becomes more valuable as the amount available increases.
- Additional dashboards, pre-built analytics, and reports become available from vendors.
- Metadata is added over time increasing the usefulness of the underlying data.
- Software tools that build and access data continually improve.

- Additional data sources can be added to the warehouse.

With such an overwhelming list of advantages, it is easy to wonder why every organization does not already have a data warehouse. The only reason is that before the availability of prebuilt warehouses, custom creation was an expensive, time-consuming, and expert-intensive process.

Thousands of organizations, including the majority of the most successful businesses in the world, have made the investment to create data warehouses. Their pioneering work has made it much easier for those starting today.

Data warehousing is now available to ordinary organizations without the vast IT resources of the giants driven by the availability of prebuilt data warehouse solutions. It should therefore be no surprise that BI is one of the fastest growing segments of the IT industry.

For those that do not yet have one, a BI system based on a data warehouse can seem like an unnecessary luxury. Once in place though, a properly built one almost instantly becomes an indispensable management tool.

Summary

The case for obtaining a BI solution based on a data warehouse has become compelling, even for businesses struggling with layoffs and drastic cost cutting. Without one it is very hard to determine how to rebuild a business model around current levels of demand.

Trying to manage a complex business in a highly challenging economic environment without a BI solution based on a data warehouse is fraught with risk. Would you set out to sea today in bad weather on a large ship without radar, GPS and a satellite radio? The fact that others had done so for many years would not make doing so a sensible choice♦.

Andrews Consulting Group (ACG) www.andrewscg.com has been helping organizations make effective use of information technology since 1984. Few IT service businesses have achieved as high a level of client satisfaction over an extended period of time as ACG. One of the most effective ways in which ACG has helped its clients is in the creation of a business intelligence solution called **RapidDecision®**.

RapidDecision® is a prebuilt data warehouse solution that follows all of the principles advocated in this paper. It therefore conforms with all the widely accepted industry best practices for data warehousing and business intelligence. A separate white paper describing **RapidDecision®** in detail is available along with much more information about it at www.rapiddecision.net. To talk directly to a **RapidDecision®** expert call (800) 775-4261 x # 5812.

ACG also publishes an on-line journal called the The JD Edwards Advisor at www.jdedwardsadvisor.com