

Avoid data management pitfalls

Various programs can be used to create management data, but discipline is required to manage it for accurate bottom-line results.

Introduction

An ongoing concern among energy companies revolves around the need for an effective strategy for management of volumetric and financial data.

Management is discovering that popular software solutions such as spreadsheets and home grown applications often contain errors which can profoundly impact the bottom line.

Unfortunately, sometimes these errors are not detected in-house, but are discovered by outside auditors and then it is “too late.”

In choosing between spreadsheets (internal systems) or commercially developed and supported systems, there are pros and cons to each. While some favor spreadsheets and internally developed database applications for company use, others struggle with the effectiveness of these non-commercial applications.

On the flip side, every company has unique requirements that certain commercially developed, or “plug and play”, systems may not be able to address effectively. Therefore, companies are faced with a decision to determine where money and effort should be spent: homegrown applications or commercially developed and supported solutions.

The goal here is to explore the various causes that contribute to the many problems inherent in data management, such as the proliferation of data silos (explained below). The following highlights several of the more prevalent data management issues seen in the industry today.

Spreadsheets - ‘Pseudo’ Programmers

Spreadsheet applications allow non-technical people the ability to become ‘pseudo’ programmers, giving them the ability to manage and manipulate data. Leveraging these new found “programming” skills, more complex modeling is applied to spreadsheets, eventually resulting in distinct applications. Tasks such as booking reserves to governmental reporting are ultimately based on information from these spreadsheets, located on someone’s individual hard drive.

Some may question if this is really a problem at all. Irrefutable evidence suggests that, individually created spreadsheets are not only inefficient, they may also potentially put the company at risk. To fully appreciate the severity of the situation, we should examine both the potential risks, as well as the possible resulting damages.

The potential risks include:

- Errors in formulas
- Audit ability
- Support
- Cost

Errors in Formulas

There are numerous reasons for spreadsheet errors, including:

- i) cryptic formulas,
- ii) the need for a copy of a formula to reside in each cell,
- iii) the need to consider all cell and formula dependencies to other worksheets when making changes, all of which can lead to hidden mistakes in spreadsheets.

Dr. Raymond Panko has conducted extensive research on spreadsheet errors. His field audit studies of actual spreadsheets in use in various organizations finds that 91% of the spreadsheets audited contained errors [Journal of End User Computing's Special issue on Scaling Up End User Development Volume 10, No 2. Spring 1998, pp. 15-21]. This finding is especially alarming given that "most audits only reported the substantive errors." Furthermore, Dr. Panko's research found that error rates occurred at this high level regardless of who was doing the modeling; spreadsheet modelers with extensive experience produced just as many errors as those with little or no experience.

Given the critical importance of financial models and the widespread dependency on spreadsheets in most organizations today, the high incidence of errors is disconcerting. The ramifications of these errors can be disastrous to organizations, especially given the scrutiny on companies' financial data and recent industry regulations that have been enacted due to corporate accounting scandals.

With that said, let us look at some of the impacts of common errors made in these applications. In 2003, TransAlta, took a \$24 million right down stemming from a bidding mistake caused by a cut-and paste error in an Excel spreadsheet. Cut-and paste errors are one of the most common error occurrences found in spreadsheets. This, along with the use of misapplied absolute and relative referencing, was indicated to be the reason behind TransAlta's \$24 million loss of.

In another case, the day after the Sarbanes Oxley legislation went into effect in the United States, Carrizo Oil and Gas discovered an error in a spreadsheet which was used tracked the average number of warrants and options outstanding. The error impacted Carrizo's financial results, skewing both the total number of shares outstanding and the income per share. The corrected number of shares for the 2004 six month period was 20.8 Million, rather than 18.9 Million as previously reported. The actual diluted net income per share was \$0.19, rather than \$0.21.

Thorough examination of corporate financials remains a priority among regulators. Therefore, company executives need to increase their efforts to avoid errors of this kind of. As one CEO recently stated, "I don't look good in orange."

Data Auditing and Integrity

Another problematic area when dealing with spreadsheets involves auditing. Internal audits of spreadsheets prove difficult because most organizations lack proper checks and balances, other methods for data validation, or error

warnings inherent in the spreadsheets. Therefore, the smallest typos and errors often go undetected while causing significant damage to overall data integrity. Furthermore, the tedious nature of auditing spreadsheets often means that the audits are sporadic and incomplete which only adds to the problem instead of serving as an authentication process.

Outside auditors have an even more difficult time verifying the accuracy of spreadsheet calculations since they are not familiar with the construction and underlying formulas of each spreadsheet. For this reason, they have to analyze the entire document in order to decipher the formulas, a task that can be extraordinarily difficult. As a result, the auditing process becomes an unexpectedly large effort for both the auditor and the company.

In one example, the shares of a Pacific Northwest self-storage company fell 7.1 percent after an independent auditing firm resigned from the auditing project. Soon after, it was discovered that the company overpaid its chief executive and another investor \$700,000 each. Company representatives claimed that the overpayment occurred when the company bought out a limited partnership owned by the Chief Executive and another investor. According to company spokesperson, an employee in the company's accounting department used the wrong spreadsheet to calculate what amount the two were owed. The miscalculated overpayment was ultimately discovered and repaid to the company six months later.

Support

In some organizations, spreadsheet applications play a central role in everyday business decisions made by the company. So, what happens when the spreadsheet creator suddenly leaves the company or becomes ill for an extended period? Since most internally developed applications have very little, if any documentation, companies are sometimes forced to turn to professional IT consultants, and spending considerable time and money trying to retrace the steps involved in creation of the spreadsheet. To avoid this scenario, business processes should be well documented and be made available to others.

A problem previously touched on is the occurrence of non-technical employees functioning as pseudo programmers. Due to their familiarity and access to spreadsheets, some feel empowered to create their own spreadsheets. These applications often grow dramatically in size with the increased business complexity. The problem lies in the portability and share-ability of these files. Even email systems can become strained as large spreadsheet applications are routinely distributed over the LAN.

Silos of Data

Another problem related to spreadsheets is controlling changes and versions. Users have little control over version a spreadsheet that is placed on a network where multiple users have access to it and are constantly modifying and resaving changes. A common reaction to this issue is that most users retain a copy of the original spreadsheet on their individual hard drive. This means that there are inevitably multiple variations of the same spreadsheet floating around

the company, which adds to the confusion of which version of the spreadsheet was actually used for the final decision or purpose. This effect is known as data siloing.

Data silos occur when information management techniques are maintained and used by a single department or individual independently of others within the company, circumventing enterprise-wide use of the information. It's not uncommon for there to be several independent data silos within a company that are actually storing and managing the same information; thereby multiplying the time and effort necessary for effective data management. As stated above, these applications or spreadsheets can take on a life of their own.

However, data silos are not strictly a spreadsheet related dilemma. Data silos can exist in any situation where multiple systems, either home grown or vendor supported, are deployed within an organization. As depicted in the diagram below, the movement of data around an organization can create problems in itself. Similar to spreadsheet use, when one incorrect system is updated and corrected while the others are not, the overall quality of the data becomes highly suspect.

All of the aforementioned problems have made it incredibly difficult for the IT professionals to support these applications and the organizations that manage the data quality.

Cost

Very few companies examine the true overall cost associated in creating a

system of various spreadsheet based applications. A spreadsheet created and maintained in-house over a 3 month period has a cost-accounting value in excess of \$25,000. This amount does not include the value attributed to the potential that other resources may use the spreadsheet every day or the amount of time to disseminate the application.

Another cost that cannot be overlooked is the cost of poor decisions that may stem from bad data. There is potential to expose the company to huge costs in misallocated funds, poor business decisions, penalties, fines and fees, not to mention the cost to the company's reputation and jeopardy of shareholder trust.

Solutions

In the world of energy companies, there needs to be an increased focus on effective enterprise data management. This area has been neglected for years. Some contributing factors include:

- Bad initial data quality (garbage in, garbage out)
- Erroneous measurement data (device errors)
- Missing data (timing or integration problems)
- Human error
- Version control & data integrity issues due to overlapping disparate systems
- Overwriting data without audit trail

To adequately and thoroughly address the issues discussed, a three-pronged approach must be followed:

1. Develop a meticulous business strategy;
2. Document the work plan and system processes;
3. Deploy a comprehensive, enterprise-wide, integrated data management and validation system.

Each company must develop a sound data management strategy and use intelligent systems to manage volumetric, transactional and financial data. Systems need to have built-in data validation mechanisms to identify data errors. In the short run, this approach may entail more work, but the long term benefit will be the alleviation the issues surrounding support.

Conclusion

In closing, while users have been given the ability to create silos of data, we need to recognize the associated negative impacts and cost of that approach. One of the greatest levers that available to business today is the use of technology. With technology we are creating data that is propagated throughout organizations at an alarming rate. However, systems (and strategy) must be put in place to manage this data. The most efficient way to do this is to implement a comprehensive, enterprise wide integrated management system. An enterprise wide system can alleviate many of the data management issues facing the industry, such as:

- Audit ability
- System knowledge and support
- Data siloing
- Version control

We have the technology to create the data, but we need humans and systems and discipline to manage the data.

Author:

Cleve Hogarth
Vice President
Quorum Business Solutions, Inc.
3010 Briarpark Drive, Suite 450
Houston, TX 77042