MITSloan
Management Review

Intelligence

Finding Value in the Information Explosion

A brief discussion about effective ways in which IT and management can work together to manage the rapid growth of data and capitalize on the opportunities it presents, by Cynthia Beath, Irma Becerra-Fernandez, Jeanne Ross and James Short

www.manaraa.com

[INFORMATION TECHNOLOGY]

Finding Value in the Information Explosion

The rapid growth of data creates new opportunities for smart analytics and improved customer service — but only if IT and business management can work together.

BY CYNTHIA BEATH, IRMA BECERRA-FERNANDEZ, JEANNE ROSS AND JAMES SHORT

Using data from its ERP system, Campbell Soup identified opportunities to reduce costs.

Modern enterprises are awash with data, and in most organizations the volume of data is expanding by 35%-50% every year. Today's companies process more than 60 terabytes of information annually, about 1,000 times more than a decade ago. Moreover, most of the data companies collect, create and manage today are unstructured — located in word processing documents, spreadsheets, images and video that can't easily be retrieved or interpreted.

How well are companies managing the data explosion and capitalizing on the opportunities it presents? To what extent are they extracting value from the data at their disposal? To answer these questions, we and our colleagues from seven IT research centers studied data- and information-related activities at 26 corporations and large nonprofit organizations in the United States. (See "Related Research.") More than 80% had

annual revenues of more than $1 billion. They represented a wide range of industry sectors, including retail, healthcare, manufacturing, education and government.

Despite the buzz around "big data," most organizations in our study were focused on the challenges of storing, protecting and accessing massive amounts of data, efforts for which IT is primarily responsible. But these organizations have not spent significant resources on the business opportunities possible with such data. Our research shows that while the IT unit is extremely competent at storing and protecting data, it cannot make key decisions that turn data into increased business value. Only a few CIOs and other IT executives reported that their organizations were succeeding at generating significant business value from their data.

Within organizations, business leaders must take the lead in making better use of data.

Why the Information Explosion Is a Top Management Concern

In every organization we studied, data growth far exceeded revenue growth. The yearly increase in the volume of stored data ranged from zero for a company that was actually shrinking in every sense (revenues, employees, etc.) to 150% for a hospital group. The median data growth rate was approximately 40% per year, which essentially doubled the volume of data stored every two years. The highest growth rates were at research universities (where some research activities

produce huge data stores) and hospital systems (where radiology and other image-intensive functions and electronic medical record applications drive storage demand). We expected that this explosion of data would allow companies to put more data in the hands of decision-makers at all levels of the organization, and in some cases our interviewees reported that it has. However, at most of the organizations, generating business value from increased amounts of data is still an aspiration. To better understand the challenge, it helps to look at the different growth drivers for structured and unstructured data.

The growth of structured data. As companies invest heavily in enterprise resource

planning systems, customer relationship management systems, radio-frequency identification tags and other technologies, the amount of data associated with transactions is multiplying. For example, companies collect data on inventory levels, transportation movements and financial transactions, enabling them to communicate more accurately with customers, accelerate supply chains, manage financial risks, optimize business processes and identify new business opportunities. Ideally, transaction data are collected and stored once. In practice, many organizations have redundant applications and databases, which add to data storage costs and make data more difficult (and more costly) to access. Increased structured data can be a double-edged sword: increased granularity creates opportunities for analytics that could lead to improved business processes and better customer service, but duplicated and conflicting data can undermine service delivery and result in conflicts over whose data are more accurate. Even with vast improvements in semantic processing, which at least in theory allow people to find data wherever they are stored, a large percentage of stored data serves no useful purpose because management has not specified how it will be used: who will make what decisions or provide what services with what data.

The explosion of unstructured data. Documents, images,

videos and e-mail make up the lion's share of existing data in most organizations, and most of the growth. This growth is fed by many factors, including regulatory requirements, the widespread use of Microsoft Office for internal communications and coordination, and increasingly vigilant tracking of customer interactions. Some companies are finding that unstructured data collected through social media and better document management systems are enabling greater sharing of information and knowledge. This sharing creates opportunities for enhanced customer service, improved operations and accelerated research and development. But realizing value from unstructured data typically requires reorganizing and indexing some of the information (for example, coding physicians' notes to indicate whether a patient is a current smoker, is a past smoker, lived with smoker, etc.). In addition, organizational barriers (such as the lack of a collaborative culture) can create obstacles to accessing potentially valuable data.

Despite these challenges, some companies are developing ways to extract important new value from their data. Aetna, for example, has significantly grown its business over the past decade by greatly ex-

panding the number of market segments in which it operates, from seven to 50. The company identified the new segments by analyzing data at the individual customer level as opposed to the more aggregated employer account level, it had previously used. Campbell Soup, for its part, has used data from its ERP system to study what it really costs to produce and deliver products to market, which has led to significant cost reductions. "We have data now that we never had before," said an IT executive. Partners HealthCare, the largest healthcare provider in Massachusetts, is reusing medical record data originally collected for clinical purposes — a mix of structured and unstructured data — to dramatically speed up medical research. As former CIO John Glaser explained, "We can cut the cost of research by a factor of five, and the time required by a factor of 10. This is a big deal. And even if those [improvements] are halved, this is still a really big deal."

Still, companies that are finding increased value in "big data" are the exception rather than the rule. Most companies have not decided who should be responsible for managing data and for driving business success. Many line business managers believe that IT should take the lead; after all,

IT runs the databases and data centers. But IT units cannot manage the information explosion on their own. Indeed, IT units struggle with two big challenges: managing the operational costs and risks associated with data, and learning how to interpret the data to enhance the organization's flexibility and performance. Meeting these challenges requires close and ongoing collaboration between IT and business managers.

Collaborating to Drive Value

To maximize the business benefits of the information explosion while minimizing its costs and risks, management must clarify accountabilities. Responsibility for data stewardship can, to a large extent, be assigned to the IT unit, but business leaders must take responsibility for defining the data value proposition and delivering on it.

At the tactical level, IT can take the lead in ensuring safe, reliable, cost-effective data storage and access, and in advancing enterprise-wide understanding of storage options and cost-effective data access. We found several practices particularly valuable.

Using tiered storage to manage costs. Tiered storage practices help manage costs by organizing data storage according to access requirements. Tier 1 storage, which has the highest

RELATED RESEARCH

▶ C. Beath, I. Becerra-Fernandez, C.F. Gibson, M. Kleeman, R.L. Nolan, J. Rockart, J. Ross, J. Short, P. Tallon and M. Winograd, "Capturing Value from the Information Explosion," Industry Report 2012-4, Center for Large-Scale Data Systems, San Diego Supercomputer Center, UC San Diego, in press.

reliability, lowest latency and highest cost, is essential for mission-critical enterprise and operational data that must be available 24/7. Tier 2, built on storage area networks for structured data and network access storage for unstructured data, is less costly yet generally sufficient for storing data where latency and reliability are less critical. Tier 3 storage, the least costly (though the most expensive to access), uses optical jukeboxes and tape systems for data that are rarely accessed. The IT unit can lead in the development of archiving processes whereby data that are rarely accessed are moved to Tier 3 storage after a specified period of time. Of course, in many countries, regulations mandate data archiving and disposal policies. Many companies have learned the hard way that not only must they have policies for retaining or destroying certain data, they must strictly adhere to those policies. IT can implement these policies, but business management must establish them and fund their implementation.

Tapping into vendor expertise. When managing tiered storage falls outside the technical expertise of an organization, it can turn to a maturing vendor market. Growing numbers of companies depend on vendors to ensure the cost-effectiveness, reliability and security of their storage of data, particularly data that are not mission-critical. Of course, identifying which data

are or are not mission-critical is a business decision, not an IT decision.

Making storage costs transparent to business users. In the past, data storage costs were a small fraction of total IT costs — in fact, most business users considered storage "free." But unstructured data and data streaming have generated huge increases in storage volumes and added real costs to businesses. Exposing these costs will reduce managers' appetites for retaining non-value-adding data. However, determining whether the value of retaining data justifies the cost is a business decision, not an IT decision.

The IT unit can provide world-class data stewardship, but that won't necessarily lead to success in the marketplace. In an information economy, businesses must leverage their data to improve business processes, introduce innovative products and services and delight their customers. Little will happen until senior managers commit to three essential practices.

1 **Identify your "sacred data."** Data about customers, sales orders, inventory items, employees and so forth constitute the bulk of most business records. As companies grow and become more complex, these data start to assume different meanings and provide different opportunities for different parts of the business. To operate seam-

lessly, organizations must identify their critical transaction data and ensure their integrity, while allowing for local variation in other data. Former J.C. Penney CIO Tom Nealon noted that in retail, the "sacred data" are the purchase orders. "That's the beginning point for any merchandise order, and that's the information that we drive right through the supply chain — from the buyer right through to the fixture in the store." In healthcare, the sacred data are the patient health records. By defining its "sacred data," management clarifies how the business will operate and sets the parameters for the organization's enterprise architecture, which the IT unit can then build out.

2 **Define the workflows that will use unstructured data.** To derive business value from unstructured data, businesses need to define the workflows that create, retrieve, change and reuse documents, messages, images and other unstructured data. The great fallacy about unstructured data is that they go hand in hand with unstructured processes. That may be true for small groups, but for unstructured data to have an enterprise impact, the business processes or workflows in which unstructured data will be used need to be defined. In particular, manual or automated processes for adding "metadata" — tags that allow unstructured data to be categorized or manipulated — need to

be defined. Once the business has decided how unstructured data should be tagged and used, the IT unit should be able to find the right tools to implement those decisions.

3 **Use data to refine business processes.** To benefit from the information explosion, businesses must commit to ongoing analysis of their data to enable continuous improvement of their business processes. As noted, Aetna uses its data to create distinct market segments. When the data shows that a particular market segment is underserved, underpriced or too costly to serve, managers have learned how to respond. Deriving benefits from information is a highly iterative process. Improving business processes, customer service and products can lead to richer data, which in turn can drive further innovations and efficiencies.

Cynthia Beath is a professor emerita at the University of Texas at Austin. Irma Becerra-Fernandez is vice provost for academic affairs at Florida International University, in Miami. Jeanne Ross is director of the Center for Information Systems Research at the Massachusetts Institute of Technology. James Short is a lead scientist at the San Diego Supercomputer Center at the University of California San Diego. Comment on this article at http://sloanreview.mit.edu/x/53409, or contact the authors at smrfeedback@mit.edu.

MIT**Sloan**
Management Review

www.manaraa.com