This paper was presented at the 1995 CAUSE annual conference. It is part of the proceedings of that conference, "Realizing the Potential of Information Resources: Information, Technology, and Services--Proceedings of the 1995 CAUSE Annual Conference," pages 2-8-1 to 2-8-7. Permission to copy or disseminate all or part of this material is granted provided that the copies are not made or distributed for commercial advantage. To copy or disseminate otherwise, or to republish in any form, requires written permission from the author and CAUSE. For further information: CAUSE, 4840 Pearl East Circle, Suite 302E, Boulder, CO 80301; 303-449-4430; e-mail info@cause.colorado.edu.
Lock the Door and Throw Away the Key?
...If we cannot provide access to data, why collect it?

Richard Pickett
Data Administrator
Princeton University
Princeton NJ
pickett@princeton.edu

Darlene Quackenbush
Director
Information Technology Planning
James Madison University
Harrisonburg VA
quackedh@jmu.edu

Abstract

Megabytes, gigabytes, terabytes, and petabytes. Institutions of higher education continue to collect ever increasing amounts of data. The intriguing challenge is not so much collecting the information as in providing useful end-user access. As campus users acquire more powerful computers with easy-to-use interfaces, they have come to expect comparable ease of use and sophistication in their administrative applications, especially when it comes to gaining access to this data.

Providing such access to data raises numerous, interrelated issues: What is the impact on data integrity, security, and data administration? Such issues need to be considered, and institutions of higher education need to institute policies in order to encourage and sustain, effective use. In the end, if the information we collect is not accessible, then why collect it?
Introduction

In many cultures, storytellers have been responsible for accumulating and recalling local histories and information vital to the welfare of their audience. As the guardians of the community's information resources, storytellers have the important responsibility of passing information to the next generation. In a real sense, they are their culture's central database, retrieving and updating data as needed. This information transfer allows succeeding generations to manage their lives and their culture more effectively.

Organizations, including higher education, are becoming increasingly aware that their own "stories"; or, in this case, the data contained in various computer applications, are a strategic asset. Our organizations need the information for day-to-day operations as well as for long-term decision-making. A number of universities and colleges have been pioneers in this area, including Virginia Tech, University of Pennsylvania, Stanford, and University of Michigan. We are increasingly aware that the availability of information is essential in all aspects of the work environment.

Increasing volume of data

This data comes in new forms and delivery mechanisms including sound, picture and virtual reality presentations and from sources as close as the palm-top we are holding or as far away as a repository on the other side of the Internet. The electronic collection of information is relatively new in relationship to the age of many institutions. In the case of Princeton, we have a computerized "history" that represents less than one-tenth of our 250 year life as a university. During this relatively brief period we have accumulated more information than imaginable just 30 years ago. Since we must assume that this collection of information will only escalate, it is important to plan for the future.

In her CAUSE94 presentation, Building the 215' Century Mind, Jennifer James noted that individuals in our current information age have to assimilate 400 times more information than their Renaissance counterparts. To deal with this information overload, users have had to become data literate. Sources and uses of data must be known and understood by all who hope to cope in this age of information.

Greater technology awareness

Awareness of technology is fast becoming part of the societal norm, presenting technology managers with a clientele who is increasingly more demanding. Slow, cumbersome or unnecessarily restrictive systems are viewed with disdain. Technical decisions are open to increased scrutiny based on new levels of awareness and understanding. As a result of this increased technological awareness, new business and academic perspectives are being considered in the delivery of data. A key component of the increasing use of business process reengineering (BPR) is technology. In fact, for BPR to succeed, technology must be creatively utilized.

Impact of Distributed Information Sources
Before the advent of personal computers, the vast majority of an organization's electronic information was centrally managed, maintained, and secured. The historical mindset did not usually include distributed processing. As we all know, the advent of personal computers has permanently changed the way computer resources are distributed. When people discovered the power of computing that they could personally control there was an increasing desire to develop individual or departmental applications, often without consideration of the organization-wide impact. Traditional computing data centers were ill-equipped to manage, much less control, the proliferation of systems.

With this explosive growth in distributed computing, there was a concomitant increase in the number of isolated pockets of data. Shadow systems with financial, personnel, and student information proliferated. Duplicate information in these fragmented data stores was collected, stored, and analyzed. In many cases inconsistent information was separately supplied to decision-makers and the resulting lack of planning and coordination of distributed computing has become a hindrance rather than a facilitator of information exchange.2

The pendulum that represents the shift from totally centralized computing to fully distributed appears to be drifting back. That does not mean we are returning to a centrally controlled data center, but does represent a return to some centralized management of data -- but with improved access.

**Needs and Benefits of Strategic Planning for Data Access**

Changes in the technology environment are encouraging increased demands for data Accuracy, Access, and Accountability. Coincidentally, an even larger and inescapably complex array of pressures facing academic institutions is leading many to focus energy and resources on these "three A's" as a possible avenue to success. The result is that one of the university's most valuable assets, information, is moved up on the strategic planning agenda. What, then, does senior management want?

**Data Accuracy**

In a nutshell, what institutions want is confidence that the information asset is derived and managed to ensure its integrity and utility. Electronic information is increasingly utilized to make strategic decisions such as budget, enrollment for classes, and faculty workload. Data that is accurate, complete and concocted according to the needs of the institution is a prized

delicacy. It is best served in ample portions enriched by easy-to-use interfaces, on top of a rich complement of powerful data storage and manipulation tools and covered with a transparent glaze of safeguards to its continued quality. The appetite for complete, timely, and meaningful information from institutional data repositories seems insatiable.

Access to Information

Information has always had a very high strategic and tactical value. In order to remain competitive, and be cost-effective, organizations must develop a new paradigm regarding access to data. Institutions must realize that data is an important asset that needs to be managed along with other valuable resources.

Our new data access paradigm must reinforce that providing access is not only viewed as a very positive characteristic, but as an essential service. Access to information traditionally controlled by individual departments within organizations is often difficult. In some instances the concerns are that the information will be incorrectly interpreted, or security violated. Convincing those that currently have this stewardship that this sharing of information is important can be a challenging task.

Shared Accountability

Closely tied to the lines of authority is security. Maximizing legal and practical protection against computer theft, eavesdropping, data tampering, and system invasion while providing for statutory requirements, business demands, and easy access by authorized users demands finesse. Information security requires ongoing attention and yet has the potential of deterring efforts toward efficient and open data access. Data access models must find a fulcrum on which to balance reasonable risk and functionality if they are to succeed.

This search for an appropriate alignment of the data access model presents interesting issues especially as computing has become more distributed. The redistribution of the data begs a redistribution of authority and responsibility. As people have become technically literate and demand, rightfully, access to information, they also need to accept the responsibility for proper access control. Particularly in client/server architectures, these lines of responsibility and authority are often difficult to construct and maintain.

External Pressures

As if these internal calls for data integrity and management, improved access to information, good information security, and intelligently designed data access schemes are not enough, external pressures of great magnitude must be considered. External pressures from state/federal agencies, legislatures and individual citizens and patrons more than ever before require academic institutions to collect and process accurate information. Accountability for funds, ability to report and forecast university activity, and maintain the overall viability of the institution is all empowered by intelligent use of data.
Sample Approaches

Princeton University

At Princeton we currently have a large number of sophisticated automated systems that support the business and academic functions. These systems vary from new client/server applications to programs that are approaching 25 years of age. As with many organizations, the computer application demands of our users often outstripped the resources of the centralized software development staff and the legacy systems. As these users became more technically capable they developed satellite systems to fill some of their needs, often with their own set of information. These systems helped to answer some of the computing needs of the university but at the same time increased the fragmentation of information. This growth of distributed data stores, coupled with the lack of full data integration in the central applications, has made it difficult to obtain information.

To remedy this situation, Princeton has embarked on a plan to develop a strong "infocentric" architecture. This architecture will serve to break down the barriers that have prevented decision-makers, from students to executives, from having current information. This is being accomplished by the establishment of a data administration function and new directions in application development. There is a strong awareness on our campus that the sharing of information is an important factor in the future management of our institution. This concept is the basis for the mission of data administration, facilitating the utilization of information across organizational units.

One of the first projects to increase access to information is the creation of the Princeton Data Mall. This is the establishment of a multifunctional data warehouse, whose first customers will be the managers of academic departments. When conceptualizing this data resource we visualized a shopping mall, with a varied collection of data "stores", each with its own specialty, some overlap in product availability, ease of window shopping, extended hours of operation, but with a common location, and architecture.

James Madison University

James Madison University faced similar circumstances: aging central systems powered by non-relational, proprietary databases; an increasing number of distributed, and disjointed, spin-off systems; and no strategy to achieve long-term viability. JMU chose to act boldly to achieve greater information sharing. The current project not only replaces the Finance, Human Resources and Student Information systems, but also aims to provide an integrated source of information to which other systems can be joined. The Integrated Information System (IIS) project is in progress and will facilitate the sharing of information and provide new functionality across departments and systems.

At the foundation of this effort is a move to relational database (Oracle) storage and retrieval. To this is added a client/server processing model and easy-to-use data query and reporting tools supported by the application vendor (PeopleSoft). Most importantly, a conscientious effort is being made to craft new policies and practices which offer supportive access to information on behalf of the university.
Success Factors

Senior Management Support

Frequently it is stated that the support of the senior management of an organization is a very important aspect of any successful project. With regard to data access, this support is imperative to help break down barriers such as financial, ownership, and shared access hurdles.

User Involvement

While without doubt the support of the executives is crucial, it is the acceptance of the project by the staff that makes implementation a success. This is particularly true since the creation of new computer systems can place a heavy toil on the individuals involved.

The role of users in the management of information technology projects is increasing. A recent survey by Deloitte & Touche of Chief Information Officers indicates that a vast majority (96%) believes that users should either lead or work as a team with information professionals in managing projects. In addition, integration of user input in the design, development, and implementation is essential. This integration increases the probability that these systems will meet the business needs of the institutions. This has been particularly important in the efforts that have been undertaken at both Princeton and James Madison University.

Data Administration

The basic concepts of effective data administration are not new. Before the advent of personal computers the data was centrally located and management was relatively simple. The advent of distributed computing and client/server architecture has complicated the situation, but the overriding goal of assuring accurate corporate information remains. While many large higher education institutions have established this functionality within their organizations, it is an important issue for all colleges and universities to address.

The establishment of data administration should be a very positive experience, one that can facilitate the utilization of information across the campus. This is a shared responsibility between the technologists and users of information requiring a coordinated effort to ensure access to accurate and timely data.

New Technologies and Methods

With the ever-increasing amounts of data that will be collected in the future, it is important that we continue to scan the horizon for new tools. Moreover, we must seek to provide additional functionality using our institutional data while seeking to involve new data providers and users.

In addition, we must find new ways to provide better service and greater value to our institution through well-planned information access and use.

**Summary**

Today, and probably for the foreseeable future, higher education is under increasing pressure to maximize the effectiveness of many resources, including data. Our ability to collect information has grown faster than our ability to distribute and understand it. To avoid making business decisions based upon incomplete, or even erroneous data, it is imperative that methods be developed to properly manage and distribute this resource.

The growing trend in distributed computing will require new paradigms in the delivery of electronic information. Data located in distributed information systems need to be evaluated for organization-wide impact. These islands of data have the potential to affect the effectiveness of managing information. As we manage these data stores information technology professionals must work closely with the potential users of information. Only through a mutual partnership will organizations be able to fully utilize the electronic stores of information.

Changing technology calls for constant reeducation. However, when access is considered, the education is not limited to learning new technical operations but also an understanding of needs across the university campus and a willingness to rethink one's role in university functions. "We've always done it that way", is out.