Digital Asset Management: Elements of an Institutional Program
Final Report on the Duke/Dartmouth Project

Contents

I. Introduction: A Call to Action

An article in the *Chronicle of Higher Education* (June 23, 2006), "Lost in a Sea of Science Data," posits that "dealing with the 'data deluge,' will be among the great challenges for science in the 21st century," and that librarians will be called in to archive and create access to these "huge amounts of information," facing cultural and financial barriers in the process. Clifford Lynch, the director of the Coalition for Networked Information (CNI), suggests that creating shared archives of raw data will engage scientists and scholars, library and information technology staff as partners in developing a central administrative strategy for the long-term survival of data and other institutional assets.

Dealing with this "data deluge" is among the greatest challenges not only for science, but for research universities more generally. Developing a broad, long-range strategy for managing both the administrative records of the university and the academic output of departments and faculty is an institutional imperative. Every unit within the university is affected by the urgent need to identify, archive, and develop access to digital information. How successful we are at addressing this challenge will determine how core university activities – advancing research and the transfer of knowledge, and the business of running the university – are pursued and reshaped in the twenty-first century. Like the article cited above, much has been written about this challenge, but examples of the deliberate, strategic development of such an institution-wide repository to manage the broad array of digital assets are few.

One obstacle has been funding. While, as Lynch points out, agencies such as the National Science Foundation are accustomed to supporting experiments and not infrastructure such as archives, other agencies such as the National Institutes of Health are requiring that NIH-funded research have a sustainability strategy built in, along with plans for guaranteeing access over the long run to the results of that research. The amount of storage required for a centralized archive of the output of a university is significant, and has not generally been factored into the budgeting process.

Another obstacle has been cultural. Individual faculty are used to keeping and managing their data themselves, often on their own hard drives, sometimes with outdated technologies. The concept of a shared repository of data is relatively new, and responsibility for its creation has not yet been determined. Some faculty are uncomfortable at the prospect of placing their intellectual property into an institutional repository, although the advantages, e.g., immediate and wide availability of research results and the development of a preservation strategy, are numerous. Where online syllabi reside, and for how long, is also an open issue. We are quickly approaching a tipping point at which faculty will expect the university to meet this critical need, particularly if we can accommodate it in an easy and cost-neutral way. This shift is well underway in administrative areas.

Under the auspices of a planning grant from The Andrew W. Mellon Foundation, Duke University and Dartmouth College jointly investigated the elements of a campus-wide strategy to manage digital assets, both administrative and academic. Between January and August of 2006, teams from the two institutions worked together to define the problem, describe the scope of

these assets, identify the potential obstacles, and outline the elements of a program. The report that follows describes the process and the results, and will, we hope, have relevance not only to the two universities but to higher education in general. The focus was not on technological solutions, but rather on developing a clearer definition of the pieces that must be integrated into a plan (a "road map") in support of an institution-wide digital asset management/enterprise content management program. This is a complex problem which, we believe, will benefit from an institution-wide solution; a national solution, while attractive, is not yet on the horizon.

The Duke-Dartmouth collaboration concluded that the creation of a Provost-level steering committee, charged with the general oversight of a campus-wide effort to inventory, evaluate, preserve and make accessible the university's digital assets of both an administrative and an academic nature, should be the critical first step in the manifestation of this leadership at the senior level. This committee would be composed of campus stakeholders, including the chief research officer, the university librarian, chief information officer, records manager, dean of the faculty, and other key administrators. It will be charged to articulate, develop, and implement a digital asset management program, and it will oversee the work of subgroups and implementation teams that will establish policies and priorities, paying careful attention to the "Elements of an Institutional Program" outlined below.

This is a call for action, and a call for leadership within our universities in response to a critical need. Inaction based on concern about the potential difficulties is not an option. Indeed, the accelerating transition from paper to digital records has already caused the irreversible loss of vital historical information. The move from paper to digital is well under way, and unless we act decisively--and immediately--the first years of the twenty first century will be forever known as an era of lost history.

Besides the intellectual impact of so much lost history, from a legal standpoint the transition to digital record keeping is a ticking bomb for our institutions. As digital systems replace paper record keeping, our carefully formulated records retention programs are becoming increasingly null and void. Without a digital equivalent to the lifecycle controls we have established for paper records, each new digital record means that our legal liability increases, and our ability to conduct our work efficiently becomes more and more difficult. Increaingly, records are no longer flowing into paper repositories and record centers, and are instead flowing onto thousands of hard disks across our institutions. Simply speaking, we're losing control of our records with every passing day.

In short, as with managing paper records, only a centralized institutional approach can reduce legal liability and ensure that digital records are captured, maintained, disposed, and preserved in accordance with our needs and the changing legal environment. In the meantime, every passing month makes a campus-wide solution more difficult; and more critical.

Now is the time for leadership and action.

II. Scope and Definitions

Information technology is an integral part of the work of universities and in most instances is a strategic enabler for the whole enterprise. Nevertheless, the products of this work and these technologies—terabytes of digital information stored on magnetic and optical disks on personal, departmental and central computers—are not managed and preserved in a systematic way. To put the problem in perspective, Dartmouth uses over 37 terabytes of storage centrally for business applications and e-mail alone, and the average demand is growing at about 6 terabytes annually. By comparison, if stored as plain text, the print collection of the Library of Congress would be equivalent to about 20 terabytes of information. In truth, universities have no idea how much data exists in their various departments. Anecdotal evidence and random sampling indicate that the quantities are very, very large, ranging from the terabyte drive of research data in the drawer of one professor, to the gigabytes of saved e-mail in the in-box of a top administrator. The quantity of vital, uncataloged, undiscovered (and undiscoverable) information that exists in risky and unprotected formats and media is growing by leaps and bounds every year. With every passing year, more is sure to be lost. We have arrived late at finding a solution due in part to the continued trust in and reliance on paper as the medium of record and in part to the independent nature and decentralized culture of higher education. A comprehensive discussion of information storage needs and capacities has been complied by Peter Lyman and Hal. R. Varian, "How Much Information?" (2003). <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>

Digital data now permeates educational institutions, reaching deeply into administrative, academic, and research processes. A digital enterprise content management strategy must therefore be responsive to, accountable to, and formed around the needs of a broad range of campus constituencies, including the Library, General Counsel, Information Technology, Risk Management, Archives, Records Management, Admissions, and Public Affairs. The President, Provost, deans, and chief research and financial officers must endorse the program. The needs of the students must remain paramount. Ultimately, the stakeholders will include every member of the campus-wide academic, research, and administrative communities; a successful program will require the broadest possible accountability.

For this report, we have adopted the University of Kansas's definition of a "digital asset" and construed it to mean management of enterprise content in the broadest sense, including any digital object that satisfies one or more of the following conditions:

- Created in fulfillment of the research, teaching, or creative work of University faculty;
- Relevant to the planning, managing, operating, controlling, or auditing of administrative functions of an administrative or academic unit of the University;
- Purchased or licensed by the University in fulfillment of an academic or administrative function under a contract that permits continuing use of the content. In addition, content created as part of "campus life" is included in this definition.

III. Context

What is digital asset management?

A Digital Asset Management program (also known as "Enterprise Content Management" outside academia) is a set of policies, processes, and systems to identify, capture, store, manage, preserve, and provide access to digital assets of long-term value to the institution.

Background and scope

Digital asset management programs are an attempt to create a technical and policy environment in which digital materials that are deemed to have long-term value can be captured, stored, and managed as institutional assets. A digital asset management program applies to digital materials the same time-tested archival and records management principles that are used for preservation and management of non-digital materials (traditional archives), ensuring access to these materials beyond the life-span of the media, formats, or custodianship of their original authors or owners.

Conceived as a campus-wide collection of resources, tools, policies and technologies, a digital asset management program is available for use by all members of the university community. It serves as an access system for current use materials, a repository for inactive materials (perhaps under the auspices of a records management program), and a storage and management infrastructure for the preservation of material for the university archives. The program should be based on a collaborative, federated model, with many campus units and partner organizations participating as content contributors and users, as well as in ongoing development and improvements to the system. A digital asset management system should be capable of handling a variety of types of information and media formats, including but not limited to single instances or sets of the following:

- Administrative documents and records
- Journal articles and other scholarly papers
- Theses and dissertations
- Learning objects used in coursework
- Other text documents
- Spreadsheet files
- Image files
- Other multimedia files
- Disk images (i.e., copies of CDs or diskettes)
- Data sets
- Computer programs

The Development of Digital Asset Management at Universities

Significant work on the idea of academic digital repositories began in the mid-1990s. Early efforts focused on making available via the web electronic versions of theses and dissertations or pre-print articles in a few scientific fields.

In the early-2000s, the idea of institutional repositories gained momentum internationally through the release of several free, open-source, systems to build them, and after publication of a number of works defining the issues around developing such repositories in universities.

Through the work of these pioneering projects, international standards have emerged that define conceptual models for digital repositories. OAIS, the Open Archival Information System, provides such a model. OAIS sets forth a framework for preserving data by establishing norms for submitting, archiving, administering, and distributing data. The Open Archives Initiative protocol for metadata harvesting (OAI-PMH) defines standards for interchange of information between repositories.

Several hundred university digital repositories now exist around the world and many others are in development, and the literature is starting to reflect lessons learned from the maturing field.

Archive/Repository systems

A handful of systems have emerged as the basis for most university repositories, and each has a community of information technologists, librarians, and administrators working with university stakeholders to continue development of the tools, standards and policies for their use. The primary systems in use are:

- DSpace – developed through a partnership between MIT Libraries and Hewlett Packard, this system is now used for about 150 repositories. DSpace was designed to accommodate a wide range of digital assets, and has been used primarily to manage, preserve, and make available faculty and student digital objects of varying types. See: http://www.dspace.org/

- Eprints – developed through funding by the US National Science Foundation and the Joint Information Systems Committee in the UK, Eprints is now in use for about 200 repositories. Its focus is strictly on providing free access to article pre-prints, and is not intended to accommodate multimedia objects or other types of files, nor to address long-term preservation needs. The Eprints system is the primary system used by advocates of the Free Online Scholarship (open access) movement. See: http://www.eprints.org/software.archives/

- Fedora – developed through a partnership between the University of Virginia and Cornell, this system aims to be an all-purpose repository, managing and providing access not only to scholarly papers, but also large collections of digital images, data sets, and other types of digital media. Because of its broader scope this system is also significantly more complex than the others, and is currently used by only 30 repositories. A commercial company, VTLS, is currently developing and marketing a version of this system that includes some user-interface enhancements and provides vendor support. See: http://www.fedora.info/community

- Digital Commons – this is the only commercial and proprietary system that is currently widely used for university digital repositories. It was originally developed by the Berkeley Electronic Press, and is now marketed and supported by ProQuest/UMI. It is similar in scope and functionality to DSpace, but includes some user-interface refinements and support from a vendor. It is currently used for about 50 repositories. See: http://www.umi.com/products_umi/digitalcommons/default.shtml#repositories

IV. Elements of an Institutional Program

Structure/Introduction

A comprehensive program to manage a university's digital assets requires a formal structure that becomes integrated into the institution in a permanent way. In addition, the program must carry sufficient administrative weight within the institution to effect meaningful change.

The first step is to gain an explicit commitment and support from the highest levels of the institution, leading to the creation of a **Steering Committee** with broad oversight. Expectations for broad participation in the program should be clearly communicated to the campus community. The Steering Committee must be chaired by an individual who can rally widespread support and secure the participation of leaders across the institution.

The Committee's charge should include the authority to research, develop, and implement an enterprise-wide digital asset management program for the institution. This group will develop policies, establish priorities, and provide oversight of the implementation, as well as hiring staff, charging subgroups as needed, creating implementation teams, and assigning accountabilities to individuals and groups across the institution. Planning will focus on three areas: priorities, policies, and implementation.

Priorities

The Steering Committee must initially tackle the task of setting priorities for attention and resources. A process for assessing needs and opportunities must become a permanent part of the institution's management structure, so that as the ground shifts and needs change the institution can quickly reassess and respond. Initially, an inventory of asset-areas should be developed to serve as the basis of priority-setting.

The task of priority-setting is one of the most important challenges the Steering Committee will face, and it is likely that a "triage" approach will be required. Information considered vital to the operation of the institution, or of irreplaceable historical value, will likely receive highest priority. Security is an important and ongoing concern. Where necessary, stopgap measures may be needed to prevent data loss. Long-term goals must be established and digital preservation choices weighed against those goals. Content at risk must be identified, and opportunities (such as grants or institutional support) must be seized. Timing may be crucial, and the group should be nimble and able to act quickly.

Policies

The technologies and processes used for digital asset management will necessarily change over time, as new needs and tools emerge. A set of principles and policies for digital asset management within the university will provide continuity as well as clarity for its users and stakeholders. Some policies and principles should be developed at the outset of the program, addressing areas such as privacy and compliance with university and government rules, where there is little room for flexibility. In other areas, policies should develop over time, allowing for patterns of need and use to emerge from the community of stakeholders. Challenges and policy areas that should be addressed by a Steering Committee early in the process are described below, along with recommendations for principles to guide development.

Scope and complexity – The need for digital asset management is so broad, and the means of addressing it so complex, that some early limitation of scope is necessary to allow a digital asset management program to get off the ground without being overwhelmed. Therefore, it is important initially to **think globally, and act locally**. By this we mean that decisions made in designing and implementing specific solutions should be made while fully considering issues of scalability, institutional compatibility, future data migration, available support, and overall institutional efficiency. Not all issues can be addressed at once. Program planners should be aware of issues that will require later attention, putting in place policies and processes that will allow for flexibility in resolving them when the time is right. Initially, the scope should be limited to a few areas where the institutional need is most urgent or the gains are most obvious, and planners should develop and test technologies, processes, and policies that can then be applied to a wider array of needs.

Incentives for participation – The decentralized and independent culture of universities makes it necessary to determine how to encourage participation in a system that will most likely be voluntary, at least in the early stages of the program's development. While the digital records of university administration are clearly owned by the university (and staff can be required to deposit them in an asset management system as part of their job), the issues are not so clear with faculty, as intellectual property ownership and workflow processes are more complicated. Therefore, planners of a digital asset management program should think carefully about how to **make it clear that participation in a university digital asset management program is in the best interest of faculty**, not just of the university. Early implementers of academic institutional repositories found that the most difficult task was getting the people with the digital assets to take the time to put them in the system. Some universities took a marketing approach, attempting to convince faculty and other stakeholders repeatedly of the importance and utility of the system. Others analyzed the work processes and opinions of faculty regarding what was important to them, and tailored the digital asset management program's functionality and messages about it to meet existing needs, workflows, and incentives. We recommend the latter approach. **A digital asset management approach should be self-evidently useful to all whose participation is required for its success, and should meet their needs and save them time.** It should be marketed and actively promoted as such. If the digital asset management program is seen as a service rather than an obligation, it is more likely to develop a base of users quickly. Technologists, archivists, and administrators can then work on the back end to ensure the longevity of the materials in the system once the program is established.

Start small, make it simple and useful, and then let it grow and evolve. The alternative approach – developing a top-down system that attempts to deal with all the complex technical and administrative problems up front, but with little concern for the experience and needs of those who will put the content into the system – is likely to result in a digital asset management system that contains few digital assets.

Confidentiality and openness – In a community as diverse as a university, the confidentiality requirements surrounding materials in a digital asset management system will also be varied. Some university repositories have taken the approach that the goal of the system is primarily to provide open access to faculty research, and that all materials in the system should be accessible to the public. While this is an appropriate (and laudable) approach for much of the output of faculty, it is seldom appropriate for administrative records, or for other materials that require confidentiality and at the same time need to be preserved and managed in digital form. **Planning for a university digital asset management system should include development of policies that differentiate among a variety of use cases, and provide for different levels of access and security depending on the submitter's and university's needs for openness or confidentiality in those cases.** It should be possible for items in the system to be deposited and managed while in a private state (to an individual or work group) and then made available to broader groups or to the general public, as appropriate. Program planners should explore whether different kinds of materials should have different default states, and whether to require or encourage openness or confidentiality for particular kinds of materials.

Terms of use, stewardship, and governance – A successful digital asset management strategy will meet both immediate and long-term needs of a variety of stakeholders. The policies that govern the rights and responsibilities of the various program stakeholders must therefore find a way to **balance potentially conflicting requirements between depositors and the institution, and account for changing needs as digital assets mature through different phases of their life cycle**. For the short term, policies will need to be developed to address who is eligible (or required) to submit materials to the repository, and what kinds of materials and uses are appropriate. Over the longer term, program managers will need policy guidance in working with digital assets whose original stakeholders may no longer be available or whose needs and desires may conflict with the institutional interest. The policies will need to **account for risk management, the relative value of the materials to the depositor and institution, and intellectual property rights of each**. What rights and responsibilities do submitters of content to the digital archive have after their data is in the system? Can items be altered once they are in the system, or must an audit trail be maintained of every submission? Under what circumstances might materials be removed from the system? If a faculty member leaves the institution, can he or she take the digital assets, or do they remain the property (or at least in the custody) of the university into whose repository they were originally submitted? A governance body for the digital asset management program should be vested with the authority to interpret and execute these policies over time, and maintain an unbroken chain of trusted stewardship over the archive.

Of course, there are many other policy areas that will need to be explored and documented in the process of developing a university digital asset management system. References to work done at other universities are included in the appendices, and should be consulted during planning.

Implementation
While each college or university may have different dynamics, this section attempts to describe a general approach to implementation that will improve the chances of success in most academic environments.

The nearly universal acceptance of digital storage for academic and administrative information occurred gradually. This is in sharp contrast to the recent explosion in availability of fast, high capacity, and inexpensive magnetic and optical disks. Suddenly, faculty are placing the outputs of their entire academic careers - papers, books, and research data - on a single spinning disk drive. Departments are storing the administrative working files needed for continued operation in a single electronic device. Paper-based record keeping systems that were designed to provide institutional control of vital information are being replaced with the equivalent of a filing cabinet on every personal computer. The space challenges of managing paper, which used to drive the creation of record retention and disposition policies, have been all but eliminated in the digital realm. As a result, colleges and universities are rapidly losing control of their own information. The risks and opportunities created by this rapid digital migration have not been matched by thoughtful guidance or support at the institutional level. The changes required to reduce the risks and to reap the benefits will challenge the way faculty and departments think about the stewardship of information.

Sponsorship and the Digital Asset Management Steering Committee – The first step in establishing an enduring institutional culture of information stewardship is to secure explicit, high-level endorsement and support. A sponsor or sponsors at the level of the Provost or Executive Vice-President can facilitate a good start and to ensure ongoing support. The second step is to create a broadly based Steering Committee to manage the program's establishment and foster collaboration across the university.

Steering Committee Activities - The steering committee will need to assign tasks to short-term teams to develop some aspects of the effort. While the committee should retain the priority setting and oversight activities, the development of day-to-day operational practices and support should be tasked to functional groups such as administrative departments, the Library and information technology organizations. The identification of practices that would best support the faculty in their research and publication efforts needs to involve faculty from a variety of disciplines. Data security planning should include the IT security office. Since changing culture is difficult, business models and a funding plan must be developed in such a way that incentives to participate are created at all levels of the organization. The funding plan will likely require the involvement of leaders at the sponsor level, deans and administrative executives.

Initially, the committee will need someone to staff the program, to call the start-up teams together and to help develop the structures needed to support the efforts across the institution. This might be a special assistant to the Provost or an assistant director in the Executive Vice President's office. The ideal person would be a member of the university with contacts and credibility. The staff person or program coordinator would report to the program sponsor(s) and to the Steering Committee.

Establishing a Permanent Organizational Structure - Informed by the work of the program coordinator and start-up teams, the Steering Committee will identify a permanent organizational home for this program within the university. This will ensure ongoing operational leadership and support. One possible scenario is that the program coordinator is tapped to take on the ongoing leadership of the effort and is moved from the sponsor's office into the permanent organizational home of the program.

Commonalities and Differences Among Different Parts of the Organization – As the system is implemented, it must be responsive to the diverse requirements of distinctive units of the institution. The differences in managing material for teaching, research and administration can be dramatic. Intellectual property and access issues arise frequently on the academic side, while administrative data may need to be summarized and analyzed on an ad hoc basis by many organizations on campus.

Traditionally, researchers have maintained their own data. Course materials are only recently being placed into the stewardship of course management systems (and many faculty are still opting out). Administrative systems have often been the purview of the business offices of the university and held totally separate from the academic processes.

Commonalities and operational efficiencies occur most obviously with regard to technology. It is less expensive per unit to purchase large quantities of storage regardless of the type of material placed there. Some software tools may cross the groups (e.g. data analysis and summary tools, database tools, operating systems, security systems, etc). There is a trend to integrate research efforts into the classroom. A data repository may provide the means for students to view and analyze data while preserving the integrity of the data for the researcher. It is likely that we will see collaboration on policy and tools at the technology level at first and that the collaboration may evolve more broadly over time.

System Attributes – Attributes of an underlying technology system can influence the program's success or failure. In light of the long-term nature of the problem being addressed and the need to encourage cultural change, the scope of attributes that must be addressed is broader than what has traditionally been encountered in system design. The Steering Committee, the permanent program sponsor, and the technology leaders of the university should identify the technology characteristics needed for success. Considerations include:

- The **technology** that supports the digital asset management effort **will evolve over time**. It is important that this evolution of background technology be transparent to the end users of the system or seen as an improvement in service.
- **The institution must retain the ability to export and migrate materials to new technology as systems evolve.**
- **It must be possible to remove materials permanently from the system.** The Steering Committee will create policies and procedures that require certain materials, particularly categories of administrative records, to be removed from the system in accordance with approved retention schedules.
- **The time it takes to access materials must be perceived as reasonable.**

- **There must be appropriate tools to access, analyze or transform the materials stored.** This implies that open storage formats and methods will be preferred over closed or proprietary approaches.
- **There must be selective degrees of access provided to materials.** Some materials will be open to the world, others to selected groups or individuals, and some materials will be open to no one but the materials' creator or the system steward.

Deciding What to Store and How Long to Keep It – A key role of the Steering Committee is to guide what is in the digital asset management system. Policies must provide guidance for administrators, faculty, and other users of the system on what can be added and what cannot. Policies should address which materials must enter the repository and what may be added at the discretion of the user. It is certain that priorities will change over time. The Committee should require periodic reviews of which materials are in the system and develop processes for identifying materials to add, specifying how long they should be retained.

Measuring Success – The steering committee should require that an evaluation / assessment model be developed to gauge the success of the endeavor. Some suggested metrics include: the quantity of data stored (in both MB and document counts); percentage participation of faculty, staff and administrative departments, numbers of collaborations using shared data, data searches performed on the materials, number of successful transitions across technologies as underlying hardware and software changes, the number of paper-based systems migrated to fully or partially digital systems; instances of inaccessible materials, and average access times for the various priority levels of materials.

Who is Responsible? – A final key role of the steering committee and the sponsors is to identify the institutional custodian(s) of the materials. It is possible that a single office may be identified or that a few offices may be chosen, but it must be clear who is responsible for which materials. There may be cases in which certain faculty members or administrators retain custodianship over materials they create, but clear policies and procedures should be proposed to clarify the flow of materials and data through the university and into the digital asset management system.

V. Duke and Dartmouth: Commonalities and Differences

Independently of this project and each other, Duke and Dartmouth had begun surveying their need for digital asset management -- Dartmouth beginning with administrative systems, particularly digital records management, and Duke with the output of their interdisciplinary centers. Despite our different starting points, similarities have been more common than differences.

Within both institutions, we see varying levels of uncoordinated, and frequently isolated, attempts to create and manage digital assets. These occur in the administrative, academic, and research arenas of both institutions. Because higher education's decentralized model gives individuals a high degree of responsibility for a particular part of the educational mission, neither institution has been able to develop the infrastructure to join the administrative, academic, and research environments.

At both Duke and Dartmouth, we find ourselves in a situation that can almost be described as circular. We must develop the wide-ranging policy, technological, and cultural infrastructure to enable assets to move smoothly into an asset management system, but there is no one starting point. Each builds on the other.

We find helpful evidence that we can develop those various infrastructure elements and that our institutions will accept them. Foremost is the desire that these services be available. A wide majority of records custodians at Dartmouth stated an interest in learning how to incorporate digital records into their records management portfolios. Duke's experience shows a similar interest, with transfers of significant electronic records to the University Archives.

Nevertheless, there are significant issues to overcome. On the administrative side of both campuses, no consistent solution emerged that will allow seamless, effective workflows to move materials between offices, assess them, then move them into a management system for permanent retention. Consequently, we find that many offices still rely largely on paper records to manage workflow, even if the genesis of the record is a form filled out online. Straddling workflows like this, version control emerges as a significant issue. Over time, this issue reemerges as a loss of institutional documentation. Departments and committees create documents but often do not recognize that these records may now be the only manifestation of that work.

The lack of flexible software may contribute to the problem. In instances when electronic records may move between offices, units have developed *ad hoc* arrangements, sometimes using email in place of more collaborative software. As Dartmouth's survey noted, "the campus email system is inadequate for current digital record keeping needs." Representatives from both Duke and Dartmouth expressed frustration with the limits of proprietary enterprise systems such as Blackboard or PeopleSoft. While such systems serve the immediate needs, they generally do not permit institutions to extract data for long-term retention in a manner providing meaningful context or addressing the issues of privacy of individuals.

Additionally, all academic and research departments look to digital solutions to enable data-gathering and research. We know that researchers gather significant bodies of research materials in electronic formats. Significantly, none of the representatives from computing or the libraries from either institution had been approached about managing these assets. This argues that without a central repository, highly publicized and easily accessible, departments and research centers seek to manage their data on their own.

In our earliest meeting, we agreed that any file targeted for retention may fit into one of three distinct stages, distinguished by the use it can fulfill:

- **Primary use**: meets initial purpose; needed now and provides business continuity for operation
- **Secondary purpose**: Initial use completed, but data's additional value is apparent
- **Long-term purpose:** Data assumed to have future, albeit undetermined, use.

Through our discussions, we have recognized that there is a significant infrastructure gap standing between the primary use and secondary and long-term purposes for the data produced. That infrastructure gap needs policy, technology, and cultural change to bridge it.

Because it appears unlikely that a single system will emerge to manage and simplify digital assets' transition from primary use to long-term retention, technology and policy will have to support one another to establish that path. Cultural change must occur if our decentralized institutions are to embrace a centralized approach to digital asset management.

VI. Duke University: The Present State of Digital Asset Management

Campus administrative units, academic departments, research centers, and student groups create a wealth of electronic records daily, some of which merit permanent retention and all require some form of management. Many of the records have paper analogs, such as publications like *Duke Magazine*; others, such as websites for student groups or financial data in SAP R/3, may have no paper version. Campus publications and departmental newsletters are increasingly available in digital form only and long term plans for management and preservation do not exist. In 2016 where will we look for course catalogs from 2002 or for the 2001 version of the faculty handbook? By not managing these data now, we are creating potential gaps in the historical and administrative record of the University.

Administrative records

The campus uses a number of enterprise systems to create and manage many of the key administrative records, such as SAP R/3 for financial and personnel records, PeopleSoft for student records, and Advance for development records. Long-term preservation of information from these systems will have to address questions of identifying and culling information of enduring value from the larger database and maintaining it within a setting that attends to both context and the privacy of individuals. While more attention and planning may be needed to insure the long term preservation of the permanent records generated by these systems, the daily management, system maintenance, and data recovery plans are already in place and centrally managed.

Administrative records created outside such enterprise systems, including content ranging from the President's annual report to campus press releases to documents used for strategic planning, do not have such central management and support and their long term survival is at greater risk. Appendix 2, which includes examples of records created by public affairs, outlines issues and concerns.

Academic and research records

As with administrative records created outside enterprise systems, records generated by academic departments and interdisciplinary centers do not have central systems in place to help them manage annual reports, conference proceedings, newsletters, faculty publications, and other such electronic content that they generate, collect, and disseminate on a daily basis. In academic year 2005-2006, Library and Archives staff surveyed the websites of over 120 interdisciplinary centers and 50 academic departments and programs to get a sense of what digital assets were

publicly available on their websites. In particular, we looked for faculty articles and publications, symposium and conference proceedings, and research data sets. Here is a summary of departments and centers that had such digital assets available on their websites:

Web sites surveyed	Articles	Proceedings	Data sets	Course materials
Academic Depts. (50 surveyed)	27 (54%)	2 (4 %)	3 (6%)	8 (16%)
Centers (123 surveyed)	52 (42%)	25 (20%)	3 (2%)	16 (13%)

In addition, 15 (30%) of the academic departments publish newsletters or similar publications on their website. Of the professional schools, only the School of Law has central management of such content in place. They use the institutional repository software Eprints to manage online faculty publications and the journals the school publishes. It should be noted that this survey only covered materials published on the web and that much content not published on departmental websites will greatly benefit from more central management.

Teaching and the records of students

The majority of online course material is managed using the course management software Blackboard, although a number of faculty and departments use other tools. At present course content placed in Blackboard is available for the current semester. Following the conclusion of class, data are not shared unless instructor neglects to turn off access to course. There is no formal retention policy for this material. Appendix 4 includes a case study that outlines the issues surrounding the management of these data.

Students also create academic content outside the course management system, such as undergraduate honors papers, master's projects, multi-media digital art and, starting in 2007, electronic theses and dissertations (ETDs). For student work and other related materials, a DSpace instance is available for storing and sharing their work. The submission of ETDs is still in the pilot phase; a repository to manage and store ETDs is under development.

In addition to academic student content, student organizations and groups create websites that document their activities. In most cases this content has no analog equivalent and may be the only record of a group's functioning. The Division of Student Affairs maintains a web-based directory of all official university-recognized student groups and their websites.

Duke University Libraries university archives

Duke-Dartmouth Mellon Project
March 2, 2006

Examples of Duke digital assets as identified by University Archives

Campus publications (examples)

- Course bulletins: http://registrar.duke.edu/bulletins/inforeg/
- Student newspaper: http://www.dukechronicle.com/
- Alumni magazine: http://www.dukemagazine.duke.edu/

Duke Websites (examples)

- Office of the President: http://www.duke.edu/president/
- Capital Campaign archive: http://152.3.224.76/campaign/index.html
- Academic departments: http://www-history.aas.duke.edu/newsletters/
- Student organizations: http://osaf.studentaffairs.duke.edu/studentorgs/stuorgs/index.html

Institutional Repositories

- Law School repository for faculty and student articles and preprints.
- http://eprints.law.duke.edu
- Undergraduate student work; managed by the College of Arts & Sciences
- https://portfolio.oit.duke.edu/index.jsp
- University Archives – SNAP server maintained by Library; used to "warehouse" electronic content captured; some content has been converted to PDF and posted on University Archives' website: http://www.lib.duke.edu/archives/history/keohane-spchndx.htm

VII. Dartmouth College: The Present State of Digital Asset Management

Over the past two years Dartmouth has been actively conducting research with respect to digital records management. In late 2004, a Digital Document and Records Management Task Force group was formed. Its broad charge was to gather and assess all aspects of the management of digital records. It was also tasked with assessing current digital management practices and the institution's readiness to move in the direction of digital-only records management.

The Task Force conducted two kinds of research. The first was a survey of the over 200 record custodians on campus. These are primarily staff members who have an active working relationship with Records Management, coordinating the way records are moved from offices to Records Management. The second and more in-depth stage of the study consisted of a series of detailed interviews with departments, researchers, and faculty. To complement the broader, horizontal approach of the survey, these interviews attempted to "drill down" into the issues specific to each department or researcher. The Task Force produced a 63-page report that details its findings and recommendations.

The survey of records custodians clearly determined, on the one hand, a readiness to move to an all-digital management approach: some 87% of respondents declared that they considered their digital documents to be records. On the other hand, it was very clear that no consistent "solution" to their management was in place: 68% said there was no plan in place to manage digital records; 83% said their documents were stored in a mix of servers and individual hard drives; and 87% said they did not apply a consistent retention schedule to their digital documents. Finally, when asked if they would be interested in learning more about digital records management, 78% said yes.

The in-depth interviews covered key administrative and academic departments, as well as faculty researchers (a detailed roster of those interviewed is in Appendix B of the Dartmouth report). Hence the scope of this phase of research was academic as well as administrative records. A standard list of questions was used for all the interviews and the department or faculty member vetted the interview notes to ensure accuracy.

One important discovery is that digital record keeping is more than simply arranging digital documents into folders on storage devices or sending them off to "retirement" in Records Management. Records enable academic and business processes, and there are significant workflows, procedures, and collaborative structures that must be anticipated and addressed. At the very least, an institutional digital records management strategy must be able to export digital items from the key campus enterprise applications (such as Oracle GL, Banner, Blackboard, Advance, just to name a few) to a records management system. In addition, for workflows that do not require such a vertical application, an institutional resource must be established to enable the secure management of documents (Dartmouth's licensing of Documentum may be an initial step in this direction).

Of the eleven issues the Dartmouth study identified, these are especially important:

1) Every area of the college is facing an ongoing transition to digital record keeping technology.
2) Workflow, collaboration, and data sharing are important aspects of records management, and the resources needed to support them digitally are not emerging.
3) Digital documents and records are at risk and are not being managed properly due to their dispersal across individual workstations.
4) Existing digital systems provide no facility for documents to comply with existing record retention policies. This results in the potential for data loss, and/or legal liability for the institution.
5) Digital record keeping demands are outpacing resources and technology.
6) Administrative and academic departments are under pressure to upgrade their systems to more sophisticated and efficient digital technology.
7) Several units are moving ahead unilaterally with digital records technology.
8) The campus email system is inadequate for current digital record keeping needs, and is inappropriate as a collaboration and workflow tool.
9) The quantity of digital academic records, particularly research data, is poised to increase dramatically. This will make the task of preserving the data in accordance with grant regulations and policy much more challenging.

On the basis of this research, the Task Force made five principal recommendations:

1) *Dartmouth should articulate and implement an institutional strategy to guide our transition to a primarily digital record keeping environment.* Specifically, the Task Force recommended a "think globally, act locally" approach: to first formulate an institutional strategy or approach and then embark on its implementation with a few key, initial projects. These early efforts would provide a solid basis on which to ramp up efforts to an institutional scale.
2) *Dartmouth should initiate a discussion of digital records management at the senior administrative level.* It was clear that the nature and scope of the challenge facing the institution requires support from the highest levels of the administration.
3) *Where appropriate, Dartmouth should pursue an "all digital" approach as a core component of its global strategy.* It is clear that the future will be increasingly (and eventually exclusively) digital, and so it is vital that the institution formulate and implement strategies that will promote, facilitate and enable the transition to all-digital practices.
4) *Dartmouth should assemble a team to formulate its institutional strategy and to identify the first steps to its implementation.* The next step should be to form a group that will begin to define, formulate and articulate the tools and strategies needed by the institution to make real progress in getting control of our digital records.
5) *The need for academic repositories and digital asset management tools for curricular and research content should be an explicit part of the institution's overall records management strategy.* In both its research data as well as the ever-growing body of digital resources used in instruction, the academic side has its own set of "record keeping" needs and issues.

One final aspect from the study deserves mention. Although the Dartmouth study was focused on digital records management, it is very clear that most departments are dealing with a mix of paper and digital documents. One aspect of the institutional strategy will need to be methods to assist departments in dealing with the mix of paper and digital records.

VII. Conclusions and Opportunities for Ongoing Collaboration

The collaborative work and learning among colleagues at Duke and Dartmouth through this project have been beneficial for both institutions. Using our joint project as a point of departure, we intend to continue to share approaches with respect to implementation strategies. We will pursue the establishment of a Provost-level committee to oversee the efforts on campus. We will also collaborate on policy development, sharing best practices and training documentation. We also anticipate the possibility of preparing joint grant applications for parts of our programs.

Our work has broader applicability and will interest other institutions and plan to share our findings with a variety of communities including EDUCAUSE, CNI, ECAR and others. We can also imagine research papers and presentations being generated. We plan to publish a journal article on this work.

This process has set both Duke and Dartmouth in a positive direction as we envision and begin to implement our strategy for creating an enterprise content management system for our respective institutions. We do not underestimate the scale, scope, and complexity of this challenge and we do not underestimate the gravity of it as well. But it is not one we can ignore. Our work must begin now.

Appendix 1: References and Resources

The case for institutional repositories and digital asset management

"Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age"
by Clifford A. Lynch
ARL Bimonthly Report 226, February 2003
http://www.arl.org/newsltr/226/ir.html

Report of the Commission on Cyberinfrastructure for the Humanities & Social Sciences
American Council of Learned Societies, November 2005
http://www.acls.org/cyberinfrastructure/cyber_report.htm

"Can Electronic Scholarship Survive?"
by Deanna B. Marcum and Gerald George
Library Issues: Briefings for Faculty and Administrators, Vol. 23, No. 6, July 2003
http://www.libraryissues.com/sub/LI230006.asp

"The Case for Institutional Repositories: A SPARC Position Paper"
by Raym Crow
August 2002
http://www.arl.org/sparc/IR/ir.html

Overviews of issues related to digital asset management in universities

"Managing Digital Assets in Higher Education: An Overview of Strategic Issues"
by Donald J. Waters
ARL Bimonthly Report 244, February 2006
http://www.arl.org/newsltr/244/assets.html

"Review of Digital Repositories"
Yale University Library, August 2005
http://www.library.yale.edu/iac/documents/DR_Review_final_27Sept05.pdf

"Managing Digital Assets: Strategic Issues for Research Libraries"
Association of Research Libraries Forum, October 2005
http://www.arl.org/forum05/

Planning and policies for digital asset management in universities

"SPARC Institutional Repository Checklist & Resource Guide"
by Raym Crow
November 2002
http://www.arl.org/sparc/IR/IR_Guide.html

"Building a Business Plan for DSpace, MIT Libraries' Digital Institutional Repository"
by Mary R. Barton and Julie Harford Walker
Journal of Digital information, Volume 4, Issue 2, May 2003
http://jodi.ecs.soton.ac.uk/Articles/v04/i02/Barton/

"Understanding Faculty to Improve Content Recruitment for Institutional Repositories"
by Nancy Fried Foster and Susan Gibbons
D-Lib Magazine, January 2005
http://dlib.org/dlib/january05/foster/01foster.html

Other recent sources

MIRACLE bibliography
http://miracle.si.umich.edu/bibliography.html

"Institutional Repository Deployment in the United States as of Early 2005"
http://www.dlib.org/dlib/september05/lynch/09lynch.html

"Academic Institutional Repositories - Deployment Status in 13 Nations as of Mid 2005"
http://www.dlib.org/dlib/september05/westrienen/09westrienen.html

A blog from the University of Houston with lots of recent references and notes on institutional repositories
http://weblogs.lib.uh.edu/weblogs/scomm/institutional_repositories/

A count of ARL libraries with institutional repositories
http://www.escholarlypub.com/digitalkoans/2006/07/18/arl-institutional-repositories-version-2/

"Taming the Digital Beast" (from Campus Technology)
http://www.campus-technology.com/article.asp?id=18574&p=1

Appendix 2: Joint Case Study and System Attributes Comparison – Public Affairs

Dartmouth College: Overview and General Information
Interviewed: Rick Adams and Roland Adams
Interviewed by Wess Jolley and Jeff Horrell
Date: 03/28/2006

Public Affairs is an information intensive office, with a variety of data management needs.

From our interview it is clear that time pressures and the demands of the "24 hour news cycle" are requiring them to produce information and respond to inquiries faster and faster all the time. They are a critical operation for the College, where a single missed piece of information may have severe consequences. Therefore, gaining control of their information and being able to access it quickly and efficiently is clearly becoming a higher priority for them as they move forward into the digital age.

In a very general way, their systems can be broken down into two major categories: Document Production, and the Maintenance of Reference Files.

1) Document Production. This category involves the production of such documents as press releases, publications, newsletters, and other information, mostly for public consumption. Each of the main players (of whom there are at least three) has their own ad-hoc manual and paper based filing systems, which they characterize as being based on each person's level of technical knowledge and personal preferences. There has been little attempt at finding a way to share these resources among the group. No central document repositories have been established for digital files.

2) Reference Files. This is primarily a paper-based "clipping file" of Dartmouth mentions (clippings make up an estimated 80% of the documents), but also includes many other types of documents, including paper copies of press releases. This paper filing system is broken down into two sections: "People" and "Subject Files". Documents that concern multiple categories are photocopied and placed in multiple files.

Several primary motivators became clear as the conversation progressed.

1) Collaboration. Much of the document production work that the department engages in is a group process, but no good tools exist for managing this collaboration. A particularly strong need in this area is for Version Control. Without good tools for managing the versioning, and knowing who has worked on a particular document that has been distributed via e-mail, the team leaders have no good way to know what is the most updated and well-reviewed version of a document.

2) Digital Asset Management. Many documents that are produced by Public Affairs are assembled using a variety of digital assets (including photos, graphics, articles, etc.). The "whole package" of these assets should be retained and managed as part of the finished product, but also remain available for reuse and redeployment into other documents.

3) Security. Due to the high security of much of Public Affairs' material, a strong concern was voiced regarding keeping these documents safe and secure, especially in a highly networked environment.

4) Loss of Web Content. Some documents are being produced and published via the web, but are not being archived appropriately. New versions supplant old, and older documents disappear

5) Time Pressures. Perhaps more than almost any other office on campus, Public Affairs is forced to respond quickly to fast breaking situations. This means that they must be able to consult information resources, assemble documents, and respond to questions extremely quickly.

In summary, Public Affairs is an office which must process and manage a large quantity of information with a small staff. Although they have not examined any automated document management options, they are very aware that their current mix of paper and digital systems does not provide them with the kind of efficiency and accuracy they would like. This need will become ever more critical as the "news cycle" continues to tighten, and they are required to produce documents and information in an ever shorter time frame. They are well-poised to take advantage of tools to help them do their work better.

Duke University: Background information:
Interviewed: Cabel Smith and James Todd
Interviewed by Tim Pyatt and Tom Harkins
Date: 03/14/2006

The Duke University Archives staff have spent substantial time with campus Public Affairs (Office of News & Communications) staff discussing how they manage electronic records and have worked to capture that information in record schedules. A retention schedule exists for Public Affairs, News, and Communications records, which includes electronic records. Public affairs, news and communications offices rely almost exclusively on electronic or "born digital" records created with word processing, spreadsheet, electronic mail, website authoring, or database programs to carry out their business and activities. As electronic records are created, managed, and stored throughout their life cycle of usefulness, general strategies for identifying basic retention needs for different sets of information must be determined.

In the summer of 2004 the Office of News & Communications to transferred the bulk of their non-current paper records with archival value to the Duke University Archives and started discussion regarding the transfer of non-current electronic files with archival value. News & Communications maintains the website for Duke's president, and when President Keohane retired in 2004, they wanted to take her site "off-line" but still preserve it. Keohane was the first Duke president to have a website. The Archives acquired the files and placed them on a server dedicated for storing archival electronic records. The next significant transfer occurred in 2005 when News & Communications migrated their website to a new content management system (CMS). Rather than migrate all press releases and news stories in electronic form to the new CMS, they chose only to keep those from 2000 on "live" on the campus website and transferred those from the 1990s to Archives.

All of the News & Communications electronic records transfers received were non-current and unrestricted (open for public use). As soon as an access/delivery system can be developed, these could be made available for public use. As a test, Archives staff created PDFs of speeches by former President Keohane from her website and made them available (http://www.lib.duke.edu/archives/history/keohane-spchndx.htm). While this works for the 100+ Keohane speeches, we do not see that as a good solution for the 18,000+ press releases and news stories.

While the content and handling of the press releases, news stories, and speeches transferred largely mirrors the content and process for their analog predecessors, appraisal and transfer of the digital media files created by News & Communications have proved far more difficult. These include interviews with faculty experts, Duke-produced news stories, promotional pieces, and clips associated with a news story or press release. Third-party media files are also present, such as a public TV interviews with campus leaders and experts. The metadata for these records are problematic as standard naming conventions are not used and often no direct linkage to the story or press release exists. However the press release has an embedded link to the streaming version of the media file, giving a connection from that side of the content. Two or more versions of the media clip are often available – an uncompressed version (with a file size of over 400 MB) as well as MPEG4, Real Media, and/or QuickTime versions (file size in 40 MB range). The process for review and transfer of this content is still in the planning stages.

ATTRIBUTE	DESCRIPTION	DARTMOUTH NOTES	DUKE NOTES
Collaboration	The ability to work together to produce digital assets. Tools may be different for collaboration within an institution and across institutional bounds.	Collaboration is critical to the success of Public Affairs. Most of their news releases, reports, publications, and other documents are collaborative efforts. Staff work with sources both on and off campus, create drafts that are worked on by other staff members and eventually finalized and approved by departmental leadership. E-mail has become the default tool for collaboration, and has proved cumbersome for larger collaborative efforts. Versioning for documents traded through e-mail is especially difficult. Many documents also require the collaboration of individuals outside the Public Affairs office, such as the President, Provost, or other officials, so ad-hoc distribution and collaboration processes are necessary. The current system (e-mail, document headers, and naming conventions) does not document who collaborates; this is a capability they would like to have.	The records of Public Affairs often represent a collaboration of campus offices in their creation; this is especially true for the media files. They create and maintain the official website for Duke's president as well the main Duke website and primarily news sites. They work closely with communications officers throughout the University.

Privacy and Confidentiality	Confidentiality is how data are treated and publicized. Privacy is an expectation; confidentiality is an intention. Information that should remain private and confidential does so in accordance with internally or externally imposed guidelines.	This is also a critical need. The Public Affairs office deals with many sensitive issues and the development of press releases and publications which address these issues may go through many drafts that should remain confidential. Conferences and deliberations as to content of these resources must also remain confidential. Sources need to remain anonymous in some cases, and journalistic integrity must be preserved. In short, all systems within Public Affairs must be secure, and only those documents which represent the official, public position of the College should be made available for public consumption.	Since these records are created for public dissemination, this is not a concern for these records. All final versions are made available to the public.
Security and Identity Management	Includes the designation of who has access to information, then identifying and verifying who gains access to it.	See above. Access to Public Affairs information must be tightly controlled, based upon a security scheme that can be customized to the document and user level.	Definitely an issue for maintaining authenticity with files sometimes altered after the release date. Unclear how many staff have access to the working and final files while they are still in the custody of News & Communications.
Life Cycle	(See data life stages, above) Documents go through different phases and schedules of retention, with different management needs at each.	Most of the final products (news releases, publications) of the Public Affairs office will require long term retention. Supporting documents, work papers, drafts, and other background information should be retained and disposed on a carefully agreed upon retention schedule. The differentiation of declared records and drafts/working papers will be an important capability.	Most records would be deemed permanent, but third party media files may have a shorter life cycle.
Assessment and Appraisal	Evaluating content of a digital object to determine its place in the life cycle.	Policies and procedures must be in place to guide retention and preservation decisions. These policies must be informed by a carefully designed retention schedule, supported and approved by the institution.	Very challenging due to the wide variety of formats and lack file naming conventions and other metadata. With early pilot tests we often had to examine each file.

Data/Content Sharing	Federated access to information with appropriate controls.	Most sharing will occur within the Public Affairs office. Documents such as press releases and publications that are designed for public consumption must have a distribution path. Information access should be controlled to the user and document level.	"Current" content is live on the campus website with source files maintained locally. The Archives have acquired some records through the use of FTP and web harvest.
Version Control	Declare versions to be records; once they're in a depository, they are frozen. Version control must gather information on who created the object, who modified it, and when.	As with collaboration, version control is a critical need that remains unmet by current systems. Other than file headers or file naming conventions, there is no reliable way to track the versions of documents as they are developed. They fear what would happen if the wrong version of a press release was disseminated, as one example. In addition, as a component of the collaboration process, the ability to track and manage document versions would allow the office to roll back to previous versions when necessary, and be sure that all players are working on the proper version at the same time.	An issue with some content as described in the background information.
Rights Management	Technological means of implementing policies of copyright or other intellectual property rights.	Not fully explored. Clearly, documents produced by Dartmouth should remain and be clearly marked as copyright to the College. There may also be a concern regarding the copyrights of materials in the clipping files that the office maintains.	While Duke holds the rights to much of the content produced, third-party media files and some faculty materials will have other rights holders. For example, if images from a faculty member's new book are part of the press release, who holds the rights to that image – the faculty member, the publisher, the photographer?
Campus Leadership and Support	Buy-in from people with resources, including their willingness to put their support and resources behind the project.	As with other departments we have interviewed, the Public Affairs office feels that a campus commitment to provide the resources and leadership required will be critical to their success. They recognize that implementing a digital system will entail a significant change to their existing work processes, and will require time to make that transition.	Since Public Affairs presents the "public face" of the university, the importance of managing this content is already a priority, but a largely unfunded one at present (beyond standard server storage and back-ups).

Storage	May include format for long-term storage or resource on which formats are kept.	Their storage needs are not huge at this time. However, a scanning project to digitize their back file of clippings would be significant. The management of digital assets such as photographs is a growing storage need.	Formats to be stored and media file size are both issues; they are anxious to transfer storage responsibility for non-current content.
Custodial responsibility	Commitment of institution to gather and manage the life cycle.	During the active life cycle of the documents, custodianship remains in the hands of the Public Affairs. However, it could be open for discussion as to what historical resources should transfer to the control of the College Archivist. There appears to be some duplication of resources between the College Archives and Public Affairs, which could be explored in more detail.	Unclear when administrative value ends for Public Affairs and custodial responsibility start for Archives for electronic files. While clearer in the "paper" era, this is not as clear with e-records.
Preservation	What are an institution's responsibilities and what are the goals? Which elements should be retained and which can be let go?	They see their reference and clipping files as a resource to be maintained in perpetuity. Other documents have active life cycles, and may process into the Archives, or be disposed, as required. Documents processing to the archives will be in a multitude of formats, and will require careful planning for preservation. An additional concern is the migration of printed documents to web formats, and the discontinuing of the paper versions. In most cases Public Affairs is trying to keep both print and paper versions of important resources available.	This is particularly an issue with third party materials – should we try to preserve that content, especially since it in many cases it is unclear who is preserving it. Migration, emulation, or some combination of file preservation strategies will need to be considered.
Workflows	The routing of documents and decisions (business process); systems needed to support workflows and incentives to participate.	The long production cycle on many of the documents produced by Public Affairs requires the ability to have both pre-defined and ad-hoc workflows. These workflows should be able to include outside parties, such as printers and ultimately the distribution to journalists and other outside parties. This capability will be critical to a successful system.	Unclear at present; does not mirror what occurred with paper records.
Packaging resources in context	Issue is how to encapsulate interrelated experiences, either by themselves, in context, or both. Limited by whatever can be captured.	This issue is relevant in regard to both web sites, and publication production. DAM tools should allow the use and reuse of content. Web Sites are not archived in any way, and they have considered screen shots to document the status of the web site at a single point in time. They get regular questions about the status of the web site at previous	The lack of metadata and linkage for media files to related stories is a challenge that must be addressed. A weekly "snapshot" of the main Duke webpage is saved.

		points, which are difficult to answer under the current system. Their web pages change daily, so this is a challenge.	
Authenticity and integrity	Strategies for guaranteeing that data and content are unadulterated.	Public Affairs has significant concerns that the information they send out has the ability to be verified as authentic. For instance, there is a danger that someone could "steal a letterhead" and make a press release that appears to come from Dartmouth. So this area is key. To address this concern the office has been releasing many of their news releases via direct text within e-mail and list-servers, rather than as word processing documents and/or attachments. This allows journalists to feel that they are sure of where the document comes from, and have a path for verification.	This is already an issue before the content is transferred to Archives. Authenticity and version control can be difficult to determine as a 1999 press release may have been reused in 2003 and resaved with a 2003 date, making it impossible to determine what exactly was released in 1999. File integrity must be documented.
Promotion and training	Implies our ability to inform users about capabilities of our systems and to train them to exploit those capabilities.	Public Affairs is enthusiastic about the prospect of making the transition to digital record keeping, so promotion may not be an issue as much as training. They have identified training and familiarization as important issues that should be addressed.	Training needs to occur both ways – creator and custodian. File management, including naming conventions, maintenance of file authenticity, and standard formats must be agreed upon. When new systems are to be deployed, those with long-term custodial responsibility need to be included as part of the planning process.
Access		Particularly with their clipping and resources files, carefully designed metadata and search strategies will be key to a successful implementation. For other documents, full text searching and other discovery paths should be available.	Metadata for files were inconsistent (no standard naming conventions); press releases are organized in directories (folders) by year and by series (i.e., all of the 1999 press releases in a single directory). Within a single year there is not a consistent filing naming convention; each file must be opened to determine release and content. A partial archive of press releases and related media files are available through the campus website. Selected content is also available through the Archives website.

Dissemination		This may be an additional category to consider adding to the system attributes. Public Affairs has the unique need to disseminate information widely, both on and off campus. Dissemination mechanisms for digital documents should be addressed in their digital system.	

Appendix 3: Joint Case Study and System Attributes comparison – Technologized Class

Dartmouth College:
Assumptions. We did not interpret "technologized" as the sophistication or complexity used by course participants (e.g., reality cave, use of Final Cut Pro, etc.). Instead, we broadened the scope, interpreting it to mean any course that made use of the course management system (CMS). We feel that by including the majority of Dartmouth courses it makes this exercise more relevant to the overall context of the Duke-Dartmouth institutional strategic planning. We did conduct the analysis thinking of a course that made use of the full range of resources offered by the CMS, including discussion threads, assessments, announcements, wikis, chat, whiteboards, etc.

For the sake of completeness we decided to analyze the course both during the term it is offered as well as afterwards.

It may make sense to read the life cycle section first, as it provides a context or frame of reference for all the other grid items.

Duke University:
Winston Atkins interviewed Neal Caidin of the Center for Instructional Technology (CIT) about issues surrounding highly technologized classes at Duke.

This case study is based on Duke's use of Blackboard.

ATTRIBUTE	DESCRIPTION	DARTMOUTH NOTES	DUKE NOTES
Collaboration	The ability to work together to produce digital assets. Tools may be different for collaboration within an institution and across institutional bounds.	During the term, a course is about as collaborative as one can imagine. During the second phase, it becomes available only to the faculty member, unless he/she chooses to make it available again to students. Once in the archival phase, the only aspect of the course web site that would be at all collaborative would be the metadata associated with it.	High degree of collaboration—learning is most effective when students are actively engaged. Tools for such coursework may be synchronous (chat, virtual classroom) or asynchronous (message boards). Synchronous data may present more challenge to capture, especially if the system captures only pointers (URLs, e.g.) in place of sites pulled in during a discussion, or non-textual elements. Asynchronous data captured more easily. Incorporating the syllabus increased the chance of capturing the process of learning within the structure of

			the class and may provide context for how the various tools are used. More collaborative tools are on the market that may not integrate within the environment.
Privacy and Confidentiality	Confidentiality is how data are treated and publicized. Privacy is an expectation; confidentiality is an intention. Information that should remain private and confidential does so in accordance with internally or externally imposed guidelines.	There are significant issues for each phase of the course life cycle. Even for the archival phase: who has authority to examine the contents of a course web site, when it may contain student work, student grades, and perhaps material under copyright?	Several aspects to this characteristic: (1) Is transmission of some, or all, data encrypted and secure? Some systems (Blackboard) present a choice regarding whether or not to encrypt selectively. (2) Access to data, including class discussions, may be controlled by instructor, allowing guests or auditing students to "sit in on" the class. (3) Some interactions (turning in assignments, assigning grades and returning assignments) require privacy at the individual level. (4) Following archiving, privacy would be difficult to ensure for some aspects, such as Chat, where knowing which participant in a discussion wrote which comment might be useful information for follow-on research, or perhaps with graded assignments.
Security and Identity Management	Includes the designation of who has access to information, then identifying and verifying who gains access to it.	During the first two life cycle phases, these issues are addressed by the CMS and the general security afforded to all enterprise servers. During the third phase, the practices associated with records management would take effect	Set at two levels. Course software permits access to students enrolled in course, authenticated through a unique identifier. Instructor may permit others to "sit in on" class, though.

Life Cycle	(See data life stages, above) Documents go through different phases and schedules of retention, with different management needs at each.	Dartmouth's current approach considers a course as having a life cycle of three phases. First, the term in which the course is conducted (e.g., Spring term 2006); second, a four year period during which the course is online and available at least to the faculty member; third, an archival stage where the course site is exported in archival format and moved from the CMS server to an archive or records management server	Mostly seen as a primary use object, either supporting class or serving as the basis for a future class. Secondary and long-term purposes may include serving as an official record of course content or possibly supporting history of pedagogy.
Assessment and Appraisal	Evaluating content of a digital object to determine its place in the life cycle.	Under Dartmouth's current life-cycle approach, assessment and appraisal are not issues in the first two stages. In the third stage—archiving—these can be complicated issues, and depending on institutional policies, apply at two levels: (1) the course level, and (2) individual document/digital object level. This is further complicated by intellectual property and copyright issues.	No retention policy, although nothing is discarded. Older versions of courses may experience problems because of software upgrades.
Data/Content Sharing	Federated access to information with appropriate controls.	During the first two phases, this kind of sharing is restricted to those officially associated with a course (e.g., instructor, student, TA). During the third phase, a new policy (as yet undefined) would apply.	Following conclusion of class, data are not shared unless instructor neglects to turn off access to course—in which case, students who had taken course would continue to have access to course materials. As stated above, there is no formal retention policy for this material.
Version Control	Declare versions to be records; once they're in a depository, they are frozen. Version control must gather information on who created the object, who modified it, and when.	This is ambiguous in the context of the course web site. It could refer to the CMS and its version; or to the versions of the applications used to create the course content; or to versions of a course (e.g., Introduction to Milton or Organic Chemistry)	Data are frozen at end of class.

Rights Management	Technological means of implementing policies of copyright or other IPR.	This is a crucial issue for course content, which could consist of original content (including software, etc.) created by faculty and students, as well as third party content and software.	Primary IPR control via restricted access to course. Nevertheless, rights management is not specifically addressed. This may become more acute if we desire to capture content not created as part of the course. In addition, if a user downloads material from the site, there is no control over material. May bring forward the question of whether course material is the property of the instructor, the university, or some combination of the two.
Campus Leadership and Support	Buy-in from people with resources, including their willingness to put their support and resources behind the project.	Institutional commitment and support would be critical to the success of any attempt to implement a repository for courses. The infrastructure could be complex and represent significant costs. There are many stake-holders with a variety of needs and expectations. Institutional policies, priorities and guidelines would be essential.	Support to provide data management for course material after the course has concluded has not gotten on the radar for university administration. University and library administration have a strong commitment to supporting the underlying infrastructure enabling instructors to use the resources.
Storage	May include format for long-term storage or resource on which formats are kept.	Archive formats for the web site are of mild concern. More formidable might be the format issues associated with the course site's content: if created with applications that no longer run, recovering that content might be impossible.	Most material are stored in a database, although content created outside the courseware, then uploaded, are stored separately from database with pointers from database to appropriate content. Material used from outside the system is not captured, only pointed to. Access to information is through a proprietary interface.
Custodial responsibility	Commitment of institution to gather and manage the life cycle.	A complex issue, particularly in the archival stage of the life cycle. See Campus Leadership and Support above.	We currently have no management process in place for older items.

Preservation	What are an institution's responsibilities and what are the goals? Which elements should be retained and which can be let go?	This would be a major issue for course content. It would involve individual documents and could involve context and setting for many of the objects contained with in the course and the course itself. The nature, definition, and application of preservation would depend on what it is that is being archived in relation to courses.	Greater clarification of the elements appropriate for long-term retention must occur. In addition, we expect that proprietary formats will lead to additional preservation problems. If synchronous and asynchronous processes are to be preserved, the information preserved will become more difficult, since referenced material from the web is retained as a pointer. Capturing that outside content to accompany the course data may raise issues of how best to effect that capture and the management of intellectual property.
Workflows	The routing of documents and decisions (business process); systems needed to support workflows and incentives to participate.	During the active life of the course, workflow is part of the course management system. In the archival phase, workflow is not an issue.	Although we expect to explore long-term course retention, we assume that the current system will not remain the access point. The program depends on whether we try to capture as much of the educational process as possible. It may be important to remember that the official repository for grades is elsewhere (PeopleSoft, in our case).
Packaging resources in context	Issue is how to encapsulate interrelated experiences, either by themselves, in context, or both. Limited by whatever can be captured.	Depending on the course or particular digital object, this could be a key issue. For much of what is archived, including the course itself, context could be crucial. See Preservation above.	As noted in Storage (above) most, but not all, content is created and stored in a database. No matter whether the content is within the database or whether the database points to it, each course's content stored in a structured way and contained within a single directory.

Authenticity and integrity	Strategies for guaranteeing that data and content are unadulterated.	This would seem to be less of an issue while the course is stored in the CMS, as that system would provide services for authentication and ways of tracking who posted what. Once the course is archived it will be important to track who has had access to the materials and might have altered them. Potentially a system such as Documentum could play a role here.	There are no tools for managing this within the current software. After the course is frozen and the data captured, strategies for ensuring authenticity would need to be developed.
Promotion and training	Implies our ability to inform users about capabilities of our systems and to train them to exploit those capabilities.	Promotion would be key to a successful operation. See Campus Leadership and Support above.	A lot of effort currently goes into promoting use of courseware systems, focused on determining how instructors want to use the software. If we determined that capturing this content (or a subset of it) was important, those channels might be available. We would also face the responsibility of determining why (therefore what) we want to preserve this: a record of the class, or documentation of the 2006 version of a highly technologized pedagogical approach.
Access		Access during the life of the course is controlled by the course management system. What type of access is needed in the archival phase depends on what it archived and what uses are intended.	Access is restricted to the faculty member and students in the course during the semester; faculty member and CIT staff have access after the semester ends