Developing a Strategy for Managing Electronic Records—The Findings of the Indiana University Electronic Records Project

Philip C. Bantin

Abstract

From June 1995 through December 1997, staff from the Indiana University Archives and University Information Technology Services undertook and completed an electronic records project partially funded by the National Historical Publications and Records Commission, designed to implement and test the "Functional Requirements for Evidence in Recordkeeping" model developed at the University of Pittsburgh. In this article, the findings of the IU project are reviewed in the context of several questions project personnel addressed during the project, including 1) Does the Pitt model ask the right questions? 2) What set of activities are required to use and implement the model? 3) What are the costs associated with implementing the model? and 4) What types of skills are required to apply the methodology.

Introduction

In June 1994 Indiana University (IU) submitted an Electronic Records Project proposal to the National Historical Publications and Records Commission (NHPRC). The proposal was a direct response to the NHPRC's call for projects designed to further research and development in the area of electronic records management, and to address specifically the set of recommendations put forward by NHPRC in *Research Issues in Electronic Records.*[1]

[1] *Research Issues in Electronic Records* (Washington, D.C.: NHPRC, 1991).

Philip Bantin was co-director of the IU Electronic Records Project. Other members of the project team included Gerald Bernbom, project co-director, and Kathy Anderson, Diana Curry, and Jeannie Kellam, project analysts.

In the development of the IU proposal, four broad concepts shaped its content, goals, and methodology. First was the conviction that electronic records management research requires useable models and generalizable results. Consequently, the IU proposal placed strong emphasis on developing model procedures and methodologies that might be applied within and beyond the IU community. Second, and linked to the emphasis on developing model statements, was a belief in the value and need for case studies as a means of advancing knowledge about electronic records. Therefore, a major goal of the IU project was to develop specific methodologies and practices that would add to the profession's knowledge about electronic records management. Third, the proposal was in response to the very real and expressed need on the IU campus for a strategy to manage and preserve those electronic records which needed to be retained for long periods of time or indefinitely. More specifically, IU data administrators had expressed in numerous public forums a need for guidance and more information on appraisal, on identifying categories of metadata required to fully describe records, and on developing strategies for improving the retrieval and preservation of electronic records. Finally, the creators of the IU proposal recognized that the project must be multidisciplinary to achieve its goals. On the IU campus, as in other complex environments, many individuals are investigating issues related to the management of electronic records. Consequently, this project was designed to involve IU personnel from a variety of disciplines.

The proposal was put forth by Indiana University under the joint sponsorship of the University Archives, University Computing Services (UCS), and the University Libraries. It was endorsed and supported by Indiana University's information policy coordinating committees: the Committee of Data Stewards and the Committee on Institutional Data.[2] The co-directors of the project were the director of the University Archives and UCS's Assistant Director for Data Administration and Access.[3]

Funding for a two-year project was approved, and work began in June 1995. (A request for a six-month extension was later submitted and approved.) Initially, the primary objectives of the project were to appraise and

[2] The IU Committee of Data Stewards is comprised of university personnel who have planning and policy-level responsibilities for data within their functional area. The data stewards, as a group, are responsible for recommending policies and establishing procedures and guidelines for university-wide data administration activities. The Committee on Institutional Data, comprised of senior university officials (typically at the level of assistant vice-president, dean, or university director), is responsible for establishing overall policy and guidelines for management and access to the administrative institutional data of Indiana University.

[3] The co-directors of the project were Philip Bantin, IU archivist, and Gerald Bernbom, a data administrator in what is now known as University Information Technology Services (UITS). Since the project ended, Bernbom has moved on to a new position within UITS dealing with digital libraries and distance education. Due to a very heavy workload in this new position, Bernbom was not available to co-author this article. However, he reviewed it and endorses and supports it. Consequently, it can truly be said that the research and findings expressed in this article are the products of a multidisciplinary team effort.

describe electronic records, develop policy and procedures for access, and define standards and procedures for preserving electronic records. The primary emphasis was on establishing archival requirements for IU's electronic records. The university offices whose records were to be reviewed were in Financial Management Support and University Enrollment Services. These areas were chosen primarily because the project co-directors had worked previously with the data stewards managing electronic records in these units, and staff in these areas had expressed a willingness to work with project personnel. During the project, team members also planned to review and analyze strategies and concepts generated by other electronic records projects, including the most prominent and detailed electronic records strategy available—the University of Pittsburgh Electronic Records Project.[4]

Project personnel were very aware that they were entering new, uncharted territory, with very few road signs to guide them along the way. The project co-directors knew that the project would be a learning process and that some, hopefully minor, changes or alterations in course would be required. But as activities on the project proceeded in the first year, staff became increasingly aware that major, not merely minor, changes in the original plans and methodology would be necessary. On the positive side, activities designed to collect data on business functions and transactions were uncovering and producing a great deal of valuable information. Problems emerged, however, when project staff attempted to use this data for analysis. As the project team struggled with next steps, it became more and more evident that the staff needed to devote much more time to developing a detailed and precise methodology. Project personnel had adopted as a strategy the identification of the processes which created records, and had identified as a goal the evaluation of how records were managed. What was lacking was a clear understanding of how to connect the two activities. As the project moved forward, team members also began to recognize that the main objective was not so much the establishment of archival requirements for electronic records, but rather a much broader goal of developing or applying recordkeeping requirements, i.e., ensuring that all records needed to docu-

[4] The University of Pittsburgh project began in February 1993 and ended in 1996. By the time the IU project began in 1995, there was plenty of information available about the Pittsburgh requirements and specifications. The availability of this information and the growing prominence of the Pitt project were major factors in the selection of the Pitt model as the centerpiece of the IU project. Other prominent electronic records projects which the IU project team reviewed and monitored included the "Building Partnerships" project administered by the New York State Archives and Records Administration, <http://unix6.nysed.gov/pubs/build.htm>; the "Preservation of the Integrity of Electronic Records" project administered by the faculty at the School of Library, Archival and Information Studies, University of British Columbia, <http://www.slais.ubc.ca/users/duranti/>; and the "Philadelphia Electronic Records" project, <http://www.lis.pitt.edu/~nhprc/perp.html>. The University of British Columbia (UBC) project (principal investigator Luciana Duranti) has emerged as a major and much discussed project, which many view as the most important alternative approach to the Pittsburgh model. In truth, the IU project team never considered testing and evaluating the UBC strategy. This decision was not based on any analysis or review of the merits of the UBC project. Rather it was based largely on the lack of substantial information on the UBC project when formative decisions were being made about the objectives of the IU project.

ment business activities were being captured, described, and preserved. Related to this observation was the recognition that the project staff were evaluating the information systems that managed the records and not the records, especially not the physical records, themselves.

These discoveries resulted in some major revisions in project objectives. One change was a significantly greater emphasis on the Pittsburgh set of functional requirements and metadata specifications. Increasingly, project staff recognized that the underlying principles and objectives of the IU project intersected with or matched those of the Pitt project. Additionally, the Pitt model offered project team members a most valuable tool—a set of requirements and specifications for evaluating IU's systems. Consequently, instead of being a source of information and guidance, the Pitt model became the centerpiece of the IU project. In essence, the primary goal of the IU project became one of applying and field testing the Pitt requirements and specifications. The other primary revision was to elevate the creation of a methodology for applying the Pitt model to a place of prominence in the project. Indeed, by the end of the project, team members acknowledged that the development of this methodology might emerge as the project's most important contribution to the profession.

In its final form, therefore, the Indiana University Electronic Records Project evolved into a project designed to implement and test the "Functional Requirements for Evidence in Recordkeeping" model developed by David Bearman, Richard Cox, and the project personnel associated with the University of Pittsburgh Electronic Records Project. IU project personnel identified four distinct stages of development for the project.

> **Stage 1—Application**: Develop a methodology for applying the "Functional Requirements for Evidence in Recordkeeping" to IU information systems.
> **Stage 2—Evaluation**: Review and evaluate IU information systems in terms of the "Functional Requirements for Evidence in Recordkeeping" and the "Metadata Specifications Derived from the Functional Requirements" developed at the University of Pittsburgh.
> **Stage 3—Recommendation**: Develop and submit a set of recommendations designed to improve the performance of the system as a recordkeeping system.
> **Stage 4—Analysis**: Review the project's application and implementation of the "Functional Requirements for Evidence in Recordkeeping" and the "Metadata Specifications." And, more broadly, critique the methodology developed in the IU project, including its use of the Pittsburgh models, in terms of effectiveness, cost, user acceptance, and skills required to implement.

For much of the project, staff had also planned to discuss recommendations with data stewards and to develop policies, standards, and software modifications designed to implement recommended changes to information systems. Unfortunately these plans were never realized. Resistance or lack of

support or even apathy were not the problem; staff simply ran out time to complete this phase of the project.

As a test site for the Pitt model, IU project personnel were attempting to address the following questions:

1. Do the underlying principles of the Pitt model stressing redefinitions of provenance and the objectives of records management provide an effective strategy for managing electronic records?
2. Does the Pitt model ask the right questions? Were all the necessary requirements and specifications present in the model?
3. Does the Pitt model include requirements and specifications that are not essential to the functional areas the project were examining?
4. What set of activities and what type of methodology are required to use and implement the Pitt requirements as a means of reviewing and evaluating information systems?
5. What are the costs associated with implementing the model? Is it cost effective?
6. What types of skills are required to apply this methodology? Who would be involved developing the methodology, and what roles would they play?
7. Will data stewards and managers understand and perceive value in this methodology and model?
8. What did project staff learn about the systems they analyzed?[5]

Summary of the Goals and Underlying Principles of the Pitt Project

An understanding of the nature of the University of Pittsburgh model and the basic underlying principles supporting it is essential to understanding the IU research. The ultimate objective of the University of Pittsburgh Electronic Records Project was to develop a statement of requirements for ensuring the preservation of evidence in recordkeeping.[6] The model consists of the following elements: the literary warrant, the functional requirements, a set of production rules, and metadata specifications. The IU project was concerned with testing two of these products: the functional requirements—the "twenty properties which are identified in law, regulation, and best practices throughout society as the fundamental properties" of evidential records;[7] and

[5] All the products of the IU Electronic Records Project can be found on the project's homepage at <http://www.indiana.edu/~libarche/index.html>.

[6] The homepage for the Pittsburgh Electronic Records project can be found at <http://www.lis.pitt.edu/~nhprc/>.

[7] David Bearman, "Item Level Control and Electronic Recordkeeping," *Archives and Museum Informatics* 10, No. 3 (1996): 207.

the metadata specifications "designed to satisfy the functional requirements for evidence," and to "guarantee that the data object will be usable over time, be accessible by its creator, and have properties required to be fully trustworthy as evidence and for purposes of executing business."[8]

At the core of the University of Pittsburgh model are redefinitions or, in one case, a more precise definition, of the archival principle of provenance, of records, and of the goals of records management. Traditional definitions equate provenance with the office of origin, resulting in an appraisal and descriptive system that focuses on administrative structure and organizational setting. But Pitt project personnel, following a trend that has been ongoing since the 1980s, redefined provenance in terms of the functions and transactions that generate a record.[9] Redefining provenance along functional lines has led Pitt personnel and a number of other archivists to recommend that the profession adopt a more precise definition of records.[10] The follow-

[8] Bearman, "Item Level Control and Electronic Recordkeeping," 208, and from introductory comments to the metadata specifications which can be found on the Pittsburgh Electronics Records Project homepage.

[9] One of the most influential, earliest endorsements of a functional approach was David Bearman and Richard Lytle, "The Power of the Principle of Provenance," *Archivaria* 21 (Winter 1985–86): 14–27. One of the earliest applications of functional appraisal was Joan K. Haas, Helen Willa Samuels, and Barbara Trippel Simmons, *Appraising the Records of Modern Science and Technology: A Guide* (Cambridge, Mass.: MIT, 1985). This methodology was later applied within a university setting in Helen Willa Samuels, *Varsity Letters: Documenting Modern Colleges and Universities* (Metuchen, N.J.: Society of American Archivists and Scarecrow Press, 1992); to the documentation of high-technology companies in Bruce Bruemmer and Sheldon Hochheiser, *The High-Technology Company: A Historical Research and Archive Guide* (Minneapolis: Charles Babbage Institute, Center for the History of Information Process, University of Minnesota, 1989), and for health care institutions in Joan D. Krizack, *Documentation Planning for the U.S. Health Care System* (Baltimore: The Johns Hopkins Press, 1994). For articles advocating a functional approach for electronic records see especially Terry Cook, "Electronic Records, Paper Minds: The Revolution in Information Management and Archives in the Post-Custodial and Post-Modernist Era," *Archives and Manuscripts* 22 (November 1994): 300–328; Terry Cook, "What is Past is Prologue: A History of Archival Ideas Since 1898, and the Future Paradigm Shift," *Archivaria* 43 (Spring 1997): 17–63; Margaret Hedstrom, "Descriptive Practices for Electronic Records: Deciding What is Essential and Imagining What is Possible," *Archivaria* 36 (Autumn 1993): 53–63; David Bearman, "Archival Strategies," *American Archivist* 58 (Fall 1995): 380–413; David Bearman, "Diplomatics, Weberian Bureaucracy, and the Management of Electronic Records in Europe and America," in David Bearman, *Electronic Evidence: Strategies for Managing Records in Contemporary Organizations* (Pittsburgh: Archives and Museum Informatics, 1994), 261–66; Greg O'Shea, "The Medium is NOT the Message: Appraisal of Electronic Records by the Australian Archives," *Archives and Manuscripts* 22 (May 1994): 68–93; and the Australian Archives homepage, "Keeping Electronic Records" at <http://www.naa.gov.au/govserv/techpub/elecrecd/keepingER.html>.

[10] For discussions of the evolution of the concept of the record and redefinitions of the term see Richard Cox, "The Record: Is it Evolving?" *The Records and Retrieval Report* 10 (March 1994): 1–16; Richard Cox, "The Record in the Information Age: A Progress Report on Research," *The Records and Retrieval Report* 12 (January 1996): 1–16; David Roberts, "Defining Electronic Records, Documents and Data," *Archives and Manuscripts* 22 (May 1994): 14–26; Glenda Ackland, "Managing the Record Rather than the Relic," *Archives and Manuscripts* 20 (1992): 57–63; Sue McKemmish, "Are Records Ever Actual?" in *The Records Continuum, Ian Maclean and Australian Archives First Fifty Years*, edited by Sue McKemmish and Michael Piggott (Clayton, Victoria: Ancora Press, 1994), 187–203; David Bearman, "Managing Electronic Mail," in *Electronic Evidence*, 188–91; Bearman, "Item Level Control and Electronic Recordkeeping," 211–14; Charles Dollar, *Archival Theory and Information Technologies: The Impact of Information Technologies on Archival Principles and Methods* (Macerata, Italy: University of Macerata, 1992), 45–48; and the Australian Archives homepage, "Keeping Electronic Records" (see footnote 9 for the URL).

ing definition of records provided by David Bearman has been gaining widespread support: Records are evidence of business transactions that document organizational functions and provide accountability. In other words, archives collect evidence, not data or information. And what is evidence? Bearman defines evidence as residing in the conjunction of data (i.e., "the record of the words, numbers, images and sounds actually made by the creator"), structure (i.e., "the relationships among these data as employed by the record creator to convey meaning"), and the context (i.e., "the relationship between the record and the activity out of which it arose").[11] If any one of these attributes is missing, the result is data, or a nonrecord. A second point to emphasize in this definition is that a record is not just a collection of data but the product of a transaction.

Another underlying principle of the Pitt model is a revised definition of the objectives of records management. Instead of focusing on the records produced by business units or on data content, Pitt project personnel argue that it makes much more sense to focus the management process on evaluating and managing the *recordkeeping systems* throughout their life cycle. This point cannot be overemphasized: the Pitt model and the IU methodology to implement the model are designed to evaluate the processes that created the record and the systems that maintain it. They are not about establishing requirements for appraising records or for modeling and defining the data content of electronic records.[12]

IU Project Results and Findings

The set of questions posed at the beginning of this article can be used as a framework to review project findings.

1. Do the underlying principles of the Pitt model stressing a redefinition of provenance and the objectives of records management provide an effective strategy for managing electronic records?

The IU project team completed their work with a very positive feeling about the value of the principles and concepts supporting the Pitt model.

[11] This definition can be found in several of Bearman's writings, but see especially "Archival Principles and the Electronic Office," in *Electronic Evidence*, 147; and "Item Level Control and Electronic Recordkeeping," 212–14.

[12] For descriptions of the recordkeeping systems concept see David Bearman, "Recordkeeping Systems," in *Electronic Evidence*, 34–70. The Australian archival community has been particularly active in defining and refining the recordkeeping concept. See Sue McKemmish, "Recordkeeping, Accountability and Continuity: The Australian Reality," in *Archival Documents: Providing Accountability Through Recordkeeping*, edited by Sue McKemmish and Frank Upward (Melbourne: Ancora Press, 1993), 9–26; Greg O'Shea and David Roberts, "Living in Digital World: Reorganizing the Electronic and Post-Custodial Realities," *Archives and Manuscripts* 24 (November 1996): 286–311; the Australian Archives homepage, "Keeping Electronic Records" (see footnote 9 for the URL); and the Australian Archives homepage, "Managing Electronic Records–A Shared Responsibility" at <http://www.naa.gov.au/govserv/techpub/manelrec/managinger.html>.

After working with this model for two and a half years and with traditional records management practices for many more years, there is no question in the minds of project personnel that the approach to managing electronic records used in the IU project is far superior to standard records management methodology. This can be demonstrated by a comparison of the effectiveness of these two strategies in terms of the following issues or activities: identification of records, identification of the universe of records necessary to document a business function, identification of how records are created, identification of documentation needed to adequately describe an event, and the overall management of records.

Identification of Records

The methodology employed by the IU project was to let functions and transactions lead staff to the identification of record-creating events. Rarely did project staff actually view data or forms or physical objects in the process of identifying records. This approach worked very well. When project personnel compared the records that were presently being created by the information system with the conceptual model of the record-creating events, they consistently discovered that the model had uncovered all the key records. The alternative to this approach is to employ traditional strategies for identifying records, which would have meant trying to identify and physically review the record or record series within the information system. However, the experience of IU project members suggests that this strategy would not be effective. Within an automated environment where records often have no fixed physical form and exist as logical, often virtual, entities, it is simply not realistic to employ a management strategy which in essence requires the physical examination and review of records. In such an environment it makes much more sense to develop an approach which de-emphasizes the need to see and handle records in order to make decisions, and which seeks and discovers the evidence needed for decision making in more accessible sources.[13]

Identification of the Universe of Records Necessary to Document a Business Function

In the identification of the records needed to document a given function, the approach of IU project staff was not to review what presently was being created, but rather to let the analysis of business functions indicate which documentation was necessary. This approach proved to be successful. In all the field tests, the conceptual model not only identified all records

[13] For a good discussion of these issues see Dollar, *Archival Theory and Information Technologies*, 35–48.

presently being created but also suggested areas where more documentation was needed. Traditional strategies, on the other hand, typically attempt to identify the universe of documentation by surveying records and asking creators to describe what they produce. This strategy is flawed because it places too much emphasis on reviewing and evaluating only those records presently being created. By so doing, such an approach may cause archivists to lose sight of the fact that one of the profession's primary goals should be to determine what kind of information is needed or required as evidence, even if this information is not presently being created.[14]

Identification of How Records are Created

In attempting to understand how offices interacted and how records were created, IU project methodology again focused on analyzing and describing business functions. The result was a very accurate depiction of how records are generated and of the flow of information in the business areas analyzed. Project personnel discovered, for example, that for several key business transactions more than one office and information system were involved. This information was critical in understanding how to document and appraise this record-creating event. Traditional records management methodology, on the other hand, typically begins with the identification of record creators and proceeds to an analysis of the records each unit generates. The experiences of the IU project team working with both types of analysis would suggest that by segmenting and compartmentalizing the record-creation process, this traditional methodology does not facilitate as accurate or comprehensive an understanding of the flow of information and of the interaction between record-creating units as does the functional approach.

Identification of the Documentation Needed to Adequately Describe the Event

In the IU project, staff focused on analyzing whether a set of specific and observable metadata was present for each transactional record. Once the Pitt metadata specifications were evaluated and selected, the actual application process went very smoothly. If the metadata existed, project staff generally had very little trouble identifying them. If they did not exist, staff were able to say precisely which items were needed. Traditional methodology would have addressed this task by attempting to create this documentation by means of observing arrangement schemes, conducting content analysis, and reviewing documentation for contextual information. There are two reasons why the products of this analysis would be far less satisfactory than those

[14] For a discussion of this concept see Samuels, *Varsity Letters*, 1–17.

generated by the methodology employed in the IU project. In the first place, there is the reoccurring problem of trying to gather information from records or systems that are simply not suited to browsing. Secondly, traditional records management methodology has not yet created a detailed and precise working definition of what constitutes a complete and comprehensive record. As a result, it would be extremely difficult to determine what is missing and what needs to be added to create a fully documented record. One of the most attractive features of the Pitt approach to documenting records is that the process more accurately reflects the way automated systems are designed to work. As a result, it is a methodology which is easier to apply and which produces higher quality products. Furthermore, because the approach more accurately reflects how data are described in systems, there is the potential for the creation of self-documenting systems.[15]

Overall Management of Records

In the Pitt model, the primary strategy for managing records is to create and maintain recordkeeping systems that capture, manage, and preserve records according to some well-defined requirements. There is the understanding that if the system that manages the records is deemed to be sound, again according to some well-defined specifications, then the records maintained within that system will also be sound. After two and a half years working with this approach, IU project personnel have come to concur with that judgment and that strategy. Project staff prefer the Pitt strategy to traditional records management practices that focus on managing the record series and the record throughout their life cycle, with the product of the analysis being the records schedule. In the future one can imagine a number of electronic records management strategies based on traditional methodologies, including linking or mapping retention schedules of paper records to electronic systems and/or attempting to locate, identify, and review major record series or the major "documentary forms"—transcripts, class schedules, balance sheets, account statements, etc.—within information systems. Based on experience working with both types of management systems, however, the IU project team believes the traditional records management strategies developed for paper records will not be effective in creating, identifying, documenting, and preserving accurate and comprehensive electronic records. On the other hand, a strategy based on the Pitt model has a good chance of achieving these goals. In the opinion of project team members, the most effective electronic records management programs the archival profession creates will include as their ultimate objectives the identification, description,

[15] For a discussion of a self-documenting system see Bearman, "Item Level Control and Electronic Recordkeeping," 213–14.

and appraisal of the functions and transactions which produce the record and the analysis of the information systems which create, manage, and preserve these records.[16]

2. Does the Pitt model ask the right questions? Were all the necessary requirements and specifications present in the model?

With one exception—the issue of policies, procedures, and metadata relating to the migration of records—IU project staff found that the Pitt model included all the requirements they thought were needed. Within the requirement for "Records: Maintained," the Pitt document addresses the issues of records being inviolate, exportable, and removable. But it does not specifically address the need for establishing specific policies and procedures for ensuring the regular migration of records or for creating metadata which provides evidence as to when and how this migration occurred and who was responsible for its implementation. IU project staff determined this was a critical issue and so added a requirement for migration both to the "Records: Maintained" section of the functional requirements and the "Disposition" section of the metadata specifications.

3. Does the Pitt model include requirements and specifications which are not essential to the functional areas the project was examining?

The IU project team felt they had a responsibility to review and consider all of the functional requirements and metadata specifications created in the final version of the Pitt model. Project personnel were not necessarily committed, however, to retaining all the requirements and specifications if there were good reasons for revising or eliminating any of them. In fact, Pitt project personnel never advocated a full-scale application of the functional require-

[16] Critics of the functional model and the emphasis on records as evidence and sources of accountability have argued that this strategy does not identify and retain the records essential for secondary research. IU project personnel support those archivists who argue that an emphasis on documenting functions and providing evidence does not exclude the retention of records which support broader historical, societal, and cultural uses. Several archivists have suggested how this might be achieved. Terry Cook suggests that it is not a "search for research value per se, but rather the articulation of the most important societal structures, functions, records creators, and records-creating processes, and their interaction, which together form a comprehensive reflection of human experience." Terry Cook, "Mind over Matter: Towards a New Theory of Archival Appraisal," in *The Archival Imagination: Essays in Honour of Hugh A. Taylor,* edited by Barbara L. Craig (Ottawa: Association of Canadian Archivists, 1992), 41. Cook has labeled this strategy "macro-appraisal," which he defines as an approach "of mirroring societal values through the functions of the record creator that focuses research instead on records creators rather than directly on society, on the assumption that those creators, and those citizens and organizations with whom they interact, indirectly represent the collective functioning of society." Cook, "What is Past is Prologue," 31. Angelika Menne-Haritz, on the other hand, argues that documentation or acquisition strategies designed to document society or to provide a true image of society by means of ranking creators or functions will not be effective. Rather she argues that the goal or the means of making the decision-making processes evident is to "make evidence accessible . . . enable the evidence to be laid open and . . . give all users the chance to interpret the evidence in their own way, giving others the chance to follow their own arguments or interpret the sources differently." Consequently, Menne-Haritz argues that "evidence is an aim, not a tool, for archival appraisal." Angelika Menne-Haritz, "Appraisal or Documentation: Can We Appraise Archives by Selecting Content," *American Archivist* 57 (Summer 1994): 541.

ments or the metadata specifications for the vast majority of information systems. What the Pittsburgh project staff did recommend was that the evaluation of the usefulness or value of a functional requirement or metadata specification be based largely on an assessment of the risk of excluding any specific elements of the model for that particular institution and that specific information system.[17] IU project staff agreed with this emphasis on risk assessment as a determining factor in selecting specific requirements or specifications. In addition, IU project personnel attempted to factor in the costs of implementing the Pitt model. From the beginning of the project, IU staff were very aware of the potential liabilities of a model that includes twenty functional requirements and over sixty metadata specifications. Such an extensive model increases the cost of the implementation, and, as one noted commentator on electronic records has observed, it may well set the project apart from other metadata initiatives. The danger is that by asking for more documentation, the goal of preserving evidence in recordkeeping might be perceived as a special and expensive project that cannot be folded into the standard procedures undertaken whenever a system is created or modified.[18] Of course, this is not how the IU project staff wanted the requirements for evidence to be perceived. Consequently, project personnel made a real effort to limit metadata specifications while at the same time ensuring that adequate evidence was retained. In sum, the position of the IU staff was that while all the Pittsburgh functional requirements and metadata specifications likely had some value, not all would be of equal value in the functional areas being tested. There would be some requirements that would clearly emerge as less important than others, would impose a minimum risk to the records in the system if excluded, and could be left out for the sake of a more cost effective model.

As indicated in the introduction to this article, the guiding principles of the IU project were to create useable models and generalizable results, and to develop model procedures and methodologies that might be applied within and beyond the IU community. Are the set of functional requirements

[17] David Bearman describes risks as including "failure to locate evidence that an organization did something it was supposed to have done under contract or according to regulation; inability to find information that is critical for current decision making; loss of proof of ownership, obligations owed and due, or liabilities; failure to document whether it behaved according to its own policies or in adherence to law; inability to locate in the proper context information which would be incriminating in one context but innocent in another." David Bearman, "Archival Data Management to Achieve Organizational Accountability for Electronic Records" in *Electronic Evidence*, 24. Helen Samuels and Tim McGovern at MIT have also developed an electronic records management strategy based on risk assessment. In a paper on the topic, they wrote "Risks are particularly great when employees in the organization do not recognize that records are, or should be created, as a consequence of transactions." Helen Samuels and Tim McGovern, unpublished paper, "Managing Electronic Evidence: A Risk Management Perspective," (1996).

[18] Margaret Hedstrom made this observation in a presentation on "Research Issues in Migration and Long-Term Preservation" at a Working Meeting on Electronic Records Research held in Pittsburgh in May 1997.

and metadata specifications selected for the IU project a model set which can be applied by IU or other institutions when reviewing and analyzing any information system? Much as the IU project team would like to answer this question in the affirmative, experience and the recommendations of other researchers in this field would strongly suggest that this is not a viable strategy. If risk assessment is the major factor in determining whether to include a particular requirement or specification, then it follows that decisions made for one specific information system in a particular institutional environment may not, indeed likely will not, be applicable to another system in another environment. In other words, the set of functional requirements and metadata specifications IU project personnel selected for the financial aid system cannot automatically be applied to records in the student records system. The assessment of risks in the two systems will be quite different, and consequently so will the set of requirements.[19]

Is it possible to put forward models of functional requirements and metadata specifications for the same types of systems operating in similar institutional environments, such as student record systems within a university setting? To date, this issue has not been tested, so IU project staff can only suggest possibilities. Certainly there will be many similarities between systems generating electronic records for similar kinds of functions and transactions. However, invariably there will be some differences, perhaps significant, in the assessment of risk by each institution. Consequently, while sharing sets of requirements and specifications will likely prove beneficial, in the judgment of the IU project team, it will not eliminate the need for each institution to conduct its own evaluation of system needs and requirements. Can one speak of a "core set" of functional requirements and metadata specifications that should be included in the analysis of any information system in any institutional environment? Again, this has not been tested; but the IU project team believes research into this issue could prove beneficial. Certainly it is not unreasonable to suggest that there exists a set of requirements that are necessary to include regardless of the specific environment or risk assessment. However, identification of a "core set" should not eliminate the need for a thorough analysis of the particular and unique requirements of an information system; rather it should serve as a point of departure for this review and analysis.[20]

[19] The author wishes to thank David Bearman for reminding him of this point. In an e-mail message to the author dated 12 August 1998, Bearman wrote: "That's the whole point of the [Pitt] model—these requirements, which are dictated by being a logical consequence of the attributes that we associate with evidence, are necessarily and correctly implemented differently in different situations because they are applied to different assessments of risk from different quarters."

[20] For example, the Dublin Core Element Set identifies a core set of fifteen metadata elements which are needed to facilitate discovery of electronic resources. For more detail on the Dublin Core Metadata see the project's homepage at <http://purl.org/metadata/dublin_core/main.html>.

Functional Requirements

Staff began by looking at the various versions of the Pitt functional requirements created over the course of the project. These versions differed not so much in terms of the specified requirements (although there were some additions and deletions of requirements during the course of the Pitt project), but rather in the way the Pitt project staff described, categorized, and arranged the requirements. For the first field test, the IU project used an early version of the requirements published by the Pitt project in spring 1993, because we felt this document included the clearest and most succinctly written descriptions of the requirements. After testing this version, however, project personnel determined that for the last two field tests they would be best served by using the final version of the Pitt functional requirements document. This final version groups the requirements into several large categories, which project personnel felt made the document easier to use and understand.

In the first field test, project staff used all of the specified functional requirements identified by the Pitt project staff. On the basis of this test and from results derived from the two subsequent field tests, however, some modifications to the Pitt model were made. All of the functional requirements at the highest level were retained: "Conscientious Organization," "Accountable Recordkeeping System," "Captured Records," "Maintained Records," and "Useable Records." But project staff eliminated categories at the subrequirements level and moved some requirements to different categories. What follows is a brief description of the specific changes (Table 1) and a rationale for why they were made.[21]

In the assessment of the functional requirements based on risks and costs, none of the requirements were judged as unnecessary or lacking in value for the functional areas under review. However, some were deemed redundant because of some unnecessary overlap between the functional and metadata requirements. Over the course of three field tests, staff identified four requirements—*Auditable, Coherent, Evidential,* and *Meaningful*—which they believed dealt exclusively with the existence of appropriate metadata and could be examined when the metadata requirements were reviewed. Consequently, these requirements were eliminated from the functional requirements section. Project staff also eliminated two requirements—*Identifiable* and *Implemented*—which they felt were issues dealt with elsewhere. In the case of *Identifiable,* the only issue thought to warrant consideration was the existence

[21] For another interpretation of the value of the Pitt Functional Requirements, see the "Functional Requirements to Ensure the Creation, Maintenance, and Preservation of Electronic Records" created by the Center for Technology in Government of the New York State Archives and Records Administration. Its most recent functional requirements model contains three primary categories rather than the five advocated by the Pitt model. The New York State model can be found at <http://www.ctg.albany.edu/projects/er/ermn.html>.

Table 1. Pittsburgh Requirements

Pittsburgh Requirement	IU Project Staff Recommendation Regarding Requirement
Compliant	Retained
Accountable – *Responsible*	Retained
Accountable – *Implemented*	Eliminated
Accountable – *Consistent*	Retained but moved to the category of "Captured Records" and combined with *Accurate*
Captured Records – *Comprehensive*	Retained
Captured Records – *Identifiable*	Eliminated
Captured Records - *Complete/Accurate*	Retained but combined with the requirement *Consistent* to form a new requirement *Accurate/Consistent*
Captured Records – *Complete/Understandable*	Retained
Captured Records – *Complete/Meaningful*	Eliminated
Captured Records – *Authorized*	Retained
Maintained Records - *Preserved/Inviolate*	Retained
Maintained Records – *Preserved/Coherent*	Eliminated
Maintained Records – *Preserved/Auditable*	Eliminated
Maintained Records - *Removable*	Retained
Usable Records – *Exportable*	Retained but moved to the category of "Maintained Records"
Usable Records – *Accessible/Available*	Retained
Usable Records – *Accessible/Renderable*	Retained
Usable Records – *Accessible/Evidential*	Eliminated
Usable Records – *Redactable*	Retained

of a unique ID, and this was addressed in the metadata section. In the case of *Implemented,* the main issue was the existence of documented procedures, which staff felt was adequately addressed under the requirement *Responsible.* One requirement, *Consistent* was moved from the area of an "Accountable System" to the category of "Records: Capture" and was combined with *Accurate* to form the requirement *Accurate/Consistent* relating to the quality of data. Project staff also moved the requirement *Exportable* from the category of "Useable Records" to the category of "Maintained Records," feeling that as a migration or maintenance issue, *Exportable* fit better into this category. In the final draft of the IU version of the functional requirements document, there are thirteen requirements. The final version of the Pitt model includes twenty requirements.

Another major revision made to the Pitt functional requirements document was to rewrite the narrative statement for each requirement with the intent of making it easier to read and use. This meant stating the intent of the function more succinctly and rewriting the statements in a clearer, more straightforward narrative. Project staff also thought the explanation of each requirement would be more useful if stated as a question or a series of questions. In most cases, the questions developed were drawn directly from the longer narrative description of the requirement created by Pitt project personnel. However, in one instance (for the requirement *Accessible*) IU project

staff expanded the focus to include questions dealing with the existence of policies and procedures relating to use and access.[22]

IU project staff spent a great deal of time developing an interpretation of the requirement *Comprehensive.* Pitt project personnel created the requirement *Comprehensive* to address the issue of whether every business event was properly captured, documented, and preserved. Initially, however, IU project personnel defined the requirement as a record content issue. Consequently, it was determined that the methodology needed a step designed to verify whether the content data required to document the transaction was being captured and preserved. After all, one could satisfy every functional requirement, but if the content of the record did not accurately and completely reflect the nature of the transaction, then all was for naught. To answer this question, IU project staff created a separate step in the methodology entitled "Identify Information Content of Transactions." In the implementation of this step, staff recommended adopting and modifying conceptual data-modeling methodology designed to verify that the content of the records faithfully reflected and represented the business relevant facts resulting from the transaction. However, in the final version of the methodology, this step has been omitted for several reasons. The project team concluded that analysis at the record content level would make this methodology cost prohibitive. Even without this step, there are serious questions about the costs of implementation; but adding data modeling and content analysis would drive the costs through the roof. One might counter with the argument that costs are not the issue here; rather, what we should be asking is whether the step is necessary to obtain the desired results. To this question, project team members would respond that it *is* a necessary activity, but the team implementing this methodology should *not* be responsible for completing this step. Data modeling or the identification of data entities and attributes and the relationships between them is a skill commonly found in the IM community,[23] and this technique is a standard part of systems design. As part of the team approach to problem solving in electronic records management, IU project personnel recommend that archivists rely on the IT community to deal with the data content issues. So, what should archivists be looking for in the requirement *Comprehensive*? IU project staff recommend asking the question posed by the Pitt project personnel: Is the system producing records for all critical business transactions? To determine this, the IU project staff suggest comparing the list of record-creating events produced in the conceptual model with the list of actual records created as identified during the review of the existing in-

[22] The list of functional requirements, as well as the metadata specifications, selected by IU project staff can be found on the project homepage.

[23] Information Management (IM) is the function within Information Technology (IT) that deals with the administration or management of data and, ultimately, information throughout their life cycles. The IM community includes data administrators and directory managers.

formation system. Consequently, the analysis remains at the transactional record level and not at the record content level.

Metadata Requirements

An initial responsibility of the IU project was to review and consider all the metadata requirements as specified in the Pitt model. However, as with the functional requirements, the IU project staff's strategy was to determine whether to retain the specification for the particular information system by assessing the risks of not including the metadata item, and by evaluating whether the benefits of including the specification were worth the costs of implementing it. As stated earlier, all the Pitt requirements and metadata specifications have value; the challenge for institutions is determining by way of a conscious, systematic, and broad-based decision-making process, which items are so critical that they need be identified as a required specification for each and every record within that system.[24]

In the assessment of risks and costs/benefits in regard to the metadata specifications, project staff asked the following two categories of questions. First, can IU live without some of the metadata specifications for this particular system? Or can IU live without some of the detail within a metadata category? Project staff found metadata in every metadata layer which fit into this category. The layers where most of the metadata were eliminated were the "Handle" layer, the *Source* metadata in the "Structural" layer, and the *Transaction Context* metadata in the "Contextual" layer. The second category asked whether metadata can be identified which, though important, does not need to be identified as a required specification for each and every record, and which might be dealt with in some less costly manner? As project personnel conducted the analysis, they became aware that a number of the metadata specifications elicited responses that were repeated, often verbatim, for many transactional records within that function and maintained within that information system. This particularly applied to some of the "Structural" layer metadata, such as the *File Encoding* metadata and the *File Rendering* metadata, and to many of the metadata items in the "Terms and Conditions" and "Use History" layers. Consequently staff began asking if all these repeatable specifications needed to be considered for every record. This is not to say that the IU staff wanted to eliminate this metadata. These specifications were important, and it was determined that IU could not risk living without them. But for the sake of cutting down on costs, could there be another

[24] In the IU project, the metadata selection passed the test of being a conscious and systematic decision-making process. But since only project team members were involved in the selection process, it was not broad-based. A next step for IU is to involve a much more diverse group in the process. This group will likely consist of members of the information management community, creators of records, representatives from legal counsel and university audit, and decision support groups.

strategy for dealing with this type of metadata? Staff determined there was a better strategy, at least for the repeatable "Structural" layer metadata. The strategy devised was to include this category of repeatable structural metadata in a separate set of requirements. This set would not be used in the analysis of each transaction but rather would become a standard part of system documentation and, most importantly, a standard set of metadata documenting migration procedures.

Note that project staff chose to employ this strategy for the repeatable "Structural metadata," but not for the repeatable "Terms and Conditions" and "Use History" metadata. Experience has shown that structural metadata has been and will continue to be a prominent part of the standard documentation compiled by the IT community. Project staff do not think it will be difficult to convince IT personnel to add a few more structural metadata items to their standard documentation. On the other hand, experience has demonstrated that the IT community does not have a high regard for access and use metadata and does not traditionally include it within the standard documentation. Project personnel are not optimistic that this pattern will change, even with the encouragement of the archival and records management communities. Therefore, it was determined that for the present the collection of repeatable access and use metadata should be more tightly controlled by remaining on the list of metadata examined for each transactional record. In summary, it would have been easiest for the project to play it safe and preserve with the record everything it takes to open and read the document. But when facing the prospect of trying to sell the methodology to the IU administration and to find a more cost-effective model, project personnel were forced to make some compromises. One of these compromises was to develop a strategy whereby some of the basic structural metadata would be captured and preserved in system and migration documentation.

Specific Recommendations Relating to Metadata

The first decision was to determine which version of the Pitt metadata specifications should be adopted. Through experience working with various versions of the Pitt metadata model, project staff felt that the last version of the Pitt metadata requirements, which organized the requirements into six layers, was the most useful and easy-to-use version. What follows is a summary of the specific changes made within each metadata layer and explanations of why project staff thought the revisions were necessary.

In the "Contextual" layer, the last draft of the IU version includes four categories of metadata. In the Pitt document there are three categories and thirteen separate pieces of metadata in this layer. To derive these four categories of metadata, project staff eliminated one mandatory and three optional

Table 2. Metadata—Contextual Layer

Pittsburgh Metadata Specifications	IU Project Staff Recommendation Regarding Specification
Transaction Context-Originator-ID	Retained
Transaction Context-Recipient-ID	Retained
Transaction Context-Copy-ID	Eliminated
Transaction Context-Business-Transaction-Type	Retained
Transaction Context-Business-Transaction Procedure Reference	Eliminated
Transaction Context-Linked-Prior Transaction	Retained
Transaction Context-Action-Requested	Eliminated
Transaction Context-Recipient Specific-Configuration Data	Eliminated
Responsibility-Originating-Organization	Retained
Responsibility-Authorization	Retained
System Accountability-System Audit-Responsible, System Audit-Implemented, System Audit-Consistent	Retained the metadata related to System Audits but moved to its own category

pieces of metadata found in the Pitt specifications (see Table 2). Staff also reorganized pieces of metadata in combinations that they thought made more sense. Thus, several pieces of metadata in the Pitt version dealing with originator and initiator were combined into one category of "Actors Involved: originator/initiator, organization, system, recipient, etc." Similarly, metadata items in the Pitt model that deal with the transaction sequence and business function type but are listed with other transactional context metadata were moved to a separate category of metadata involving process activities represented by the records. Project staff also removed the metadata dealing with time and instance of the transaction from the general list of context metadata and created a separate category for this data. Project personnel believe this new arrangement better emphasizes the critical metadata involved in this layer and ultimately makes the document easier to use. Finally, a group of metadata dealing with audits was moved out of the "Contextual" layer, and a new layer entitled "Accountability" was created, which includes a reference to collecting metadata citing any system audits. Project staff believe that the existence of audit documentation is such a critical issue that it deserves to be its own layer or category.

In the "Handle" layer, project personnel felt only one piece of metadata was needed—the unique identifier which designated the data as a record—rather than the six pieces of metadata recommended in the Pitt version (see Table 3). Staff believe that the other mandatory items related primarily to the Pitt strategy of creating Metadata Encapsulated Objects (MEO's) or self-describing metadata objects. If one were not adopting that strategy, these items were judged redundant because they described issues relating to domain and time that were already dealt with in the "Context" layer.

In the "Terms and Conditions" layer, the final IU version includes two categories of metadata, as opposed to the four categories and fourteen pieces

Table 3. Metadata—Handle Layer

Pittsburgh Metadata Specifications	IU Project Staff Recommendation Regarding Specification
Record Identification-Record-Declaration	Retained
Record Identification-Transaction-Domain-Identifier	Eliminated
Record Identification-Transaction-Instance-Identifier	Eliminated
Information Discovery Content-Content-Description-Standard	Eliminated
Information Discovery Content-Content-Descriptor	Eliminated
Information Discovery Content-Record-Natural-Language	Eliminated

of metadata identified in the Pitt version (see Table 4). This was achieved in part by eliminating two mandatory and four optional metadata items specified in the Pitt model. As in the "Contextual" layer, this reduction was also a result of combining metadata items and of moving some into their own separate categories. Project staff combined all the Pitt metadata dealing with restrictions on access into one category of documentation. This same strategy was employed for metadata items dealing with conditions for access and use. Again, the reasoning was to create a document that was easier to use. It should be noted that in combining these items a description was added that defined the types of metadata sought under these categories. Project staff also determined that the disposition requirements metadata that the Pitt model included in the "Terms and Conditions" layer merited being placed in its own category dealing exclusively with metadata on disposition issues. The six metadata items dealing with disposition in the Pitt version were then reduced down to three pieces, and a fourth specification dealing with migration data was added.

Table 4. Metadata—Terms and Conditions Layer

Pittsburgh Metadata Specifications	IU Project Staff Recommendation Regarding Specification
Restrictions Status-Access-Rights-Status	Retained
Restrictions Status-Use-Rights-Status	Retained
Access Conditions-Access-Conditions-Resolver	Retained
Access Conditions-Resolver-Terms	Eliminated
Use Conditions-Use-Conditions-Resolver	Retained
Use Conditions-Use-Terms/Use-Citation	Eliminated
Use Conditions-Use-Terms/Redacted-Record-Rule	Retained
Use Conditions-Use-Terms/License-Terms	Eliminated
Disposition Requirements-Removal-Authority	Retained but moved into its own category
Disposition Requirements-Retention-Policy-Citation	Retained but moved into its own category
Disposition Requirements-Retention-Authority Issuance	Eliminated
Disposition Requirements-Retention-External-Authority	Eliminated
Disposition Requirements-Retention-Period-End-Time	Retained but moved into its own category
Disposition Requirements-Disposition-Instruction-Code	Eliminated

Table 5. Metadata—Use History Layer

Pittsburgh Metadata Specifications	IU Project Staff Recommendation Regarding Specification
Use History-Use-Type	Retained
Use History-Use-Instance-Time	Retained
Use History-Use-Instance-User	Retained
Use History-Use-Evidential Consequences	Retained

In the "Use History" layer, all the metadata identified by the Pitt project were retained (see Table 5). However, project staff decided to split one of the Pitt metadata requirements related to use into two separate pieces. Thus the IU version has five pieces of metadata as opposed to four pieces in the Pitt version. Project personnel recognize that some archivists may disagree with the decision to include all the use history metadata. The decision was motivated in part by local conditions. For the past year or so, Indiana University administrators have frequently expressed concerns about the potential misuse of "university internal" data, and this anxiety has resulted in recommendations for further restricting access to this information. It is hoped that the incorporation of additional metadata defining what, how, and by whom data was used will go a long way towards addressing these concerns and maintaining a high level of access to university records. This decision provides an example of the concept expressed earlier in this article that the inclusion of specific pieces of metadata will be dependent on the needs of the particular system under review and of the conditions existing at the specific institution.

In the "Structural" layer, the Pitt version includes six categories of metadata containing twenty-seven individual metadata items of which only four are considered optional under all conditions. In the final IU version, there are four categories of structural metadata (see Table 6). As explained above, some of the Pitt structural specifications—a total of nine metadata items—were moved to a separate list of metadata that would become a standard part of system and migration documentation. In addition, project staff eliminated eight pieces of structural metadata from the Pitt list of specifications. The structural metadata included in the IU version are designed to render the record as originally viewed; to identify the relationships, values, and meanings of files or structures that are used to represent the record; and to identify the source that created the record.

In addition to evaluating the merit of the various metadata requirements, project personnel also spent a good deal of time rewriting the requirements in an attempt to either clarify, simplify, or summarize. Some of the Pitt requirements were difficult to understand or interpret largely because of a lack of explanation on the nature of the requirement or because of unclear language. The goal was to create a more user-friendly document that could be understood by an archivist with a fair-to-average knowledge of automated systems.

Table 6. Metadata—Structural Layer

Pittsburgh Model Specifications	IU Project Staff Recommendation Regarding Specification
File Identification-File-ID	Retained
File Encoding-File-Modality, File-Data-Representation, Data-Codes, Compression-Method, Encryption-Method	Retained but moved to a separate list of metadata which would become a standard part of system and migration documentation
File Rendering-Rendering-Rules	Retained
File Rendering-Application-Dependency, Software-Environment Dependency, Hardware-Dependency, Representation-Standard	Retained but moved to a separate list of metadata which would become a standard part of system and migration documentation
Record Rendering-File-Linking-Rule, File-Interchange Standard	Eliminated
Content Structure-Content-Structure	Retained
Content Structure-Content-Data Set	Retained
Content Structure-Application-Dictionary	Retained
Content Structure-Delimiters/Labels	Eliminated
Content Structure-Data Value-Lookup Tables	Retained
Content Structure-Data View-at Creation	Eliminated
Content Structure-Version-Relationships	Retained
Content Structure-Set-Relationships	Retained
Content Structure-Dynamic-Relationships	Retained
Source-Data-Source	Retained
Source-Data-Source-System-Documentation	Eliminated
Source-Data Capture-Instrument-Type	Eliminated
Source-Data Capture-Instrument-Settings	Eliminated
Source-Source Data-Quality	Eliminated

To summarize, the metadata document the IU project created contains twenty-one metadata categories, significantly fewer than the sixty-seven metadata items included in the Pitt model. To achieve this reduction, project personnel eliminated twenty-three mandatory or optional metadata specifications from the Pitt list, and combined or reorganized numerous other metadata items. Another nine items in the Pitt "Structural" layer were moved to a separate list to be included in system documentation. In addition, the IU version differs in the way it is organized and described. Specifically, in an attempt to make the document more user friendly both in language and in organizational structure, project staff combined specific pieces of metadata under more general headings and added plain, fairly nontechnical language to describe the types of metadata that should be collected for that category.

4. What set of activities and what type of methodology are required to use and implement the Pitt requirements as a means of reviewing and evaluating information systems?

The Pittsburgh model is a "statement of requirements needed to ensure the preservation of evidence in electronic form and not the application requirements for archival or records management systems."[25] In other words,

[25] From the introduction to the "Functional Requirements for Evidence in Recordkeeping," which can be found on the Pittsburgh project homepage.

the model provides a tool for the evaluation of systems; it does not provide a methodology for applying that model. Thus a major goal of the IU project was to define the application methodology required to use and implement the Pitt requirements as a means of reviewing and evaluating information systems.

The first step in developing such a methodology was to review the professional literature. Unfortunately, not much guidance was available from these sources. Certainly there is some excellent information on functional decomposition and data modeling in the literature of the information management field, but this methodology says very little about identifying transactions. Information on the methodology used in systems analysis and design is plentiful and useful, but this literature tells us little or nothing about creating and maintaining records.[26] Regrettably, the literature in the archival and records management fields was not very helpful at all in developing the methodology. Clearly, archivists and records managers are not yet focusing on these issues; and, unfortunately, most are simply not trained to perform these functions. Therefore, much of the methodology had to be created and tested by IU project personnel.

After several false starts and modifications, the following methodology was derived. It is possible, however, that this methodology will be refined and changed as it continues to be tested. Nonetheless, IU project staff think the steps outlined here offer a viable means of reviewing information systems using a version of the requirements outlined in the University of Pittsburgh study.

Methodology—Functional Analysis

In performing functional decomposition, project staff initially addressed three basic questions: (1) what is a business function? (2) how does one gain the information needed to do the analysis? and (3) what are the products of the process?

Clearly this analysis will not succeed unless one has a very clear understanding of the concept of a business function and can articulate how a function relates to business processes, transactions, and activities. Functions, in fact all the concepts named above, relate to official actions pertaining to the business or mission of the enterprise. In the IU project a business function was broadly differentiated from a business process by emphasizing that a function describes *what* is done in the organization, independently of *who* does it. This is in contrast to a business process with its emphasis on outputs for

[26] An excellent source on functional decomposition is James Martin, *Information Engineering* (Englewood Cliffs, N.J.: Prentice Hall, 1989). An outstanding source on the methodology used in information systems analysis and design is Edward Yourdon, *Modern Structured Analysis* (Englewood Cliffs, N.J.: Yourdon Press, 1989). A useful source on data modeling is G. Lawrence Sanders, *Data Modeling* (Danvers, Mass.: Boyd & Fraser Pub. Co., 1995).

a particular customer or market, and on *how* work is done. Functions (and subfunctions) are comprised of one or more related transactions, which are themselves comprised of related activities.

In gathering this information, project personnel relied on two basic sources: system documentation and interviews with one or more staff members from the business area. Project staff found that documentation identifying functions and subfunctions existed within the target offices. The major objective then became one of verifying and refining the data, usually by adding additional subfunctions and narratives describing how this function was performed. It is difficult to generalize about whether archivists can routinely expect information about functions to be readily available. But the project team can say with some confidence that information managers understand the concept of functional decomposition very well; it is part of their process of logical data modeling. In identifying people to interview regarding business functions, project staff started with the individuals they felt understood the entire work flow process. At this point in the analysis, an overview of the business area is required, not details of individual activities. To obtain this information usually means talking with senior-level staff from a business area rather than the individuals who mange the data systems on a day-to-day basis.

What were the products of this analysis? At the most basic level, they consisted of a list of first-level functions, subdivided into more detailed second-level and, if necessary, third-level subfunctions. The general guidelines for the IU project were that even major business areas typically have only six to twelve first-level functions, and that for each of these first-level functions, there are typically between three and eight second-level subfunctions, with the lower numbers being very common. Project staff found, however, that to really understand the nature of the function, they had to go beyond a simple listing of functions and subfunctions. What had to be developed were short narrative descriptions of what occurred within the given function. These descriptions were designed to answer the following questions: 1) who was involved in executing the function? 2) what was the nature of the action? 3) how do the primary activities within the function relate to one another? and 4) what were the products of the function?

Methodology—Identification of Transactions

The process of identifying business transactions proved to be more difficult than uncovering and describing functions. There were several reasons for this, beginning with the fact that the identification of business transactions is a relatively new concept for archivists and the IM community alike.[27] Lack

[27] Within the IT community, a transaction does not relate to a business activity but rather refers to an individual and discrete automated activity, such as changing the value of a data element. Nor-

of experience manifested itself at the most basic level—in defining just what the project staff were seeking. Consequently, the first step was to develop a working definition of a transaction with which the project staff was comfortable and which they could articulate to their audiences. The working definition adopted by the project broadly defined a transaction in the following terms: 1) an official action, related to the business or mission of the university; 2) a public, not private, action involving more than one person; 3) an action undertaken, having a beginning, and 4) an action completed, having a definite end point.

But this general definition only provided a starting point in the search for transactions. The next critical signpost in the identification of transactions was the point in the analysis at which project staff began to clearly recognize that record-creating events were occurring. During the functional decomposition process, project personnel were drilling downwards through functions, subfunctions, and finally to transactions. At some point, they identified business activities that were clearly leading to some real documentation—these were the record-creating events or transactions that they sought. This was the level or point where project staff could begin the process of forming and combining the transactional units that would be documented. The key was not to settle at too high or too low a level. If project personnel began documenting the business activities too high, they would likely have missed some critical business transactions. On the other hand, if staff drilled down too low, they would have gotten bogged down in unnecessary detail, would likely have missed relationships between business activities, and ultimately would have spent an unacceptable amount of time and money on the process. One of the keys to the overall acceptance of this methodology and to its cost effectiveness will be the development of a working definition of how archivists can more accurately and effectively identify the level of the function/subfunction/transaction continuum at which the analysis of record-creating events typically occurs.

Despite efforts to formalize and describe the process, there is no question in the minds of project personnel that a certain amount of judgment will always be required in identifying a transaction. This does not mean that the transaction is an arbitrary unit created for the sake of the analysis, for the concept of the transaction is clearly grounded in actual and real functions and activities within a business area. Instead, project staff view the transaction in part as a conceptual tool used to organize descriptions of business activities in a way that discloses their meaning. Consequently, the size and boundaries of the transaction one chooses to document and analyze can legitimately be regarded as a product of one's choice and judgment. In the IU project, staff

mally, numerous automated transactions will be undertaken in the completion of a business transaction. Moreover, in generating conceptual models, data analysts typically do not decompose business functions down to the transactional level. The term business transaction is not part of their professional vocabulary.

became increasingly aware that they could and should, for the sake of a more cost-effective model, be combining as many small, related activities or transactions as practical into one larger transactional unit. Reviewing the work of the project, staff believe they could have done even more to aim their documentation efforts at a higher transactional level and still not have omitted or left undocumented any significant record-creating events.

In describing transactions, project staff considered the following: 1) what is the official action? 2) who is taking or initiating the action? 3) what objects are being acted upon? 4) what individuals are interacted with? and 5) what are the record-creating events? The rationale for creating these descriptions was both to better understand the nature of the transaction and to begin developing responses to the questions which would be asked in the evaluation stage of the review. The products of this process were a general description of each transaction and a series of short, one- or two-sentence responses to each of the categories identified above.

Methodology—Review of Existing Record Systems

Up to this point, project staff had been engaged in an exercise of defining the business transactions of the organization and determining the categories of information a recordkeeping system should be capturing, making available, and preserving as evidence of these transactions. In the review of existing record systems, the staff attempted to describe how existing information systems actually collected and managed this data.[28] The first step in the process was to identify the information systems used to support the business transactions that had been identified. This information would typically have been gathered as part of one of the earlier interviews with staff in the business units. Once the system was identified, analysis could begin. This process of review was based on existing documentation (technical, procedural, and policy-level) and on interviews with staff in the business units and in the computing organization. One product of this review process was an organized collection of documentation and field notes that described the primary data stores and data items in the system, the primary processes or processing cycles, the overall flow or movement of data, security mechanisms, and other procedures or policies governing the operation and management of the system. Another set of products included a list of the records or documentation presently produced for this business transaction and brief descriptions of how the system managed each of these major transactional

[28] If the archivist/records manager is participating in the design of a new system, then, of course, this step is not required. In developing a new system, only two steps of the IU methodology would be required: the creation of a conceptual model of business transactions and of recommendations on these transactions. Clearly, participating in the design process is much less time consuming and more efficient than reviewing and modifying existing systems. It provides further evidence that the best, most effective strategy for archivists is to be actively involved early in the information management process, preferably at the design stage.

records. Included in this description were responses to the following questions: how created? how described? (which contained information on context, terms and conditions, disposition, accountability, use, and structure), how used? how maintained? and how distributed? As one can clearly see, these questions mirror many of the concerns expressed in the Pitt functional and metadata requirements; and, of course, this was no accident. The goal of this review was to gather the types of information that would be needed at the evaluation stage.

Methodology—Evaluation of Systems

Only after all this data was collected could the staff begin to address the questions posed in the Pittsburgh functional requirements document. Project staff had now identified 1) what they were trying to document (the transaction) and 2) how the present system managed this information. Project personnel were now ready to determine to what degree the system was capturing and retaining records, i.e., evidence (comprising information on content, context, and structure) of a transaction. To achieve this goal, IU staff had only to compare or map the two sets of descriptions using the Pitt requirements and specifications as their guide or model. In other words, the evaluation process became an exercise in comparing "what is" with "what should be" at the transactional record level.

In completing the evaluation of the information systems managing the transactions under review, project staff determined that what would work best was to frame the responses in terms of specific categories of compliance. For the functional requirements analysis, the following categories were employed: 1) satisfied, 2) partially satisfied, and 3) not satisfied. For the metadata requirements, project personnel created the following categories of compliance: 1) metadata is available electronically as part of the record; 2) metadata is available in electronic or paper form, and the metadata or a citation to this metadata is linked to the record; 3) metadata is available in electronic or paper form, but the metadata or a citation to the metadata is not linked in any way to the record; and 4) metadata is not available anywhere. After the compliance category was selected, project staff prepared a brief statement explaining in more detail how the system met or did not meet the requirement.

Methodology—Recommendations Relating to System Improvements

Once the analysis of how the information system was managing the transactional records was completed, project staff met to determine what recom-

Table 7. Tasks Performed and Time Required to Complete Task

Task	Time to Complete Task
Gather and compile information on all business functions undertaken by Financial Aid	76 hours
Identify and describe six major functions and thirty-one subfunctions	20 hours
Identify and describe twenty-one business transactions	45 hours
Review existing information system—describe general documentation	30 hours
Review existing information system—describe how system manages the transaction "Processing Appeals"	60 hours
Review existing information system—describe how system manages the transaction "Processing Pell Grants"	25 hours
Evaluate results and create written analysis—for transaction "Processing Appeals"	30 hours
Evaluate results and create written analysis—for transaction "Processing Pell Grants"	10 hours
Discuss evaluation and generate written recommendations for both transactions	8 hours
Total Time to Complete Activities Listed Above	304 hours or 38 working days or approx. 7.5 working weeks

mendations would be made for improving the performance of the system. It was recognized that not all the problems identified were of equal rank or of the same level or degree of seriousness. Therefore, project personnel designed four categories or levels of recommendations: Level 1—highest priority recommendations, Level 2—concerns, Level 3—for your information, and Level 4—does not warrant any attention at this time. Each recommendation in the first three categories included an explanation of the problem, and recommendation(s) on how the unit might address it. This document along with a short cover letter outlining the process was then forwarded to the appropriate data stewards and managers, and a meeting to discuss the various issues was scheduled.

5. What are the costs associated with implementing the model? Is it cost effective?

During the project, staff maintained a record of time spent completing each task for the two field tests on transactions within the financial aid business area. The Financial Aid Information System is a very large system which processes data for an eight-campus system with over 90,000 students, approximately 60 percent of whom receive financial aid. The system consists of 27 files, 357 programs, and 69 on-line screens. Table 7 provides a breakdown of tasks performed and the time required to identify and describe all business functions and subfunctions within the business area of financial aid and to analyze and evaluate how the Financial Aid Information System managed two major transactions: processing appeals and processing Pell Grants.

While these figures are useful as a starting point in the discussion of costs, IU project personnel believe these totals are inflated and are not truly reflective of the time which will be required to complete future analyses.

There are several reasons for this belief. First, experience working with the methodology would indicate that the more one performs the various activities included in the methodology, the better one gets and the less time it takes. Secondly, many of the problems encountered when evaluating transactions within a system appear again and again. These reoccurring patterns will allow one in essence to duplicate earlier responses when completing the evaluation statement, and this of course translates into significant time savings. Another issue which contributed significantly to inflating the amount of time expended on the analysis was the considerable time spent on rewriting and revising the products of the project, especially in the early stages. During the course of the project, as staff did less reworking of the methodology and became more proficient in its use, efficiency increased dramatically. Another factor inflating the time commitment was the tendency of project staff to overdescribe in applying sections of the methodology. A general rule that project personnel followed was to add more detail whenever there was doubt about how much information to include. Part of this strategy no doubt stemmed from the fact that project staff were not quite sure yet what information was absolutely necessary to complete the evaluation. One area where project team members feel that less information is needed is in the description of how the system manages the transaction. Another area where more data than required was collected was in the descriptions of functions, subfunctions, and transactions. Overall, project personnel feel confident in asserting that the project could have achieved the same results with far less description, resulting in a correspondingly dramatic reduction in the overall time commitment. There was also the problem of having one of the project analysts come on board near the end of the project and having to learn the language of archivists and the specific methodology "on the fly" during the actual application of the methodology. This project analyst did not have the luxury of a lengthy orientation period to learn these new concepts and techniques; essentially she had to learn by trial and error. This of course added to the time required to complete the work. Finally, in evaluating the figures obtained from the field tests, it must be recognized that some of these tasks—notably the gathering of information on all business functions and the creation of general descriptions of system documentation—are one-time-only activities, the results of which can be applied to the analysis of all transactions within that business area and managed by that information system.

Factoring in all these issues, project personnel feel strongly that the time committed to completing the tasks in the two field tests is not indicative of the real time costs. Team members are confident that these figures could easily be reduced by at least one-third, and perhaps by one-half. Following through on this judgment and reducing the total time figure for completing work on the two transactions by one-third, one arrives at the following totals:

203 hours or approximately twenty-five work days or five working weeks. These figures are a more accurate indicator of the time required to perform these tasks, but even these totals may represent the high end of the scale.

It would be instructive to use these same figures, i.e., one-third less time than the field test numbers, to compute an hourly total for completing an analysis for all the major transactions for the function "Award Financial Aid." In this major function, there are six record-producing transactions—some smaller, some bigger than "Processing an Appeal." Therefore, if we estimate that it would take about the same time as "Processing an Appeal" to complete an analysis of each of the remaining transactions, the results would be as follows: total time to complete work on the function "Award Financial Aid": 566 hours or approximately seventy work days or approximately fourteen working weeks.

Using these same figures, it would be instructive to make one further projection. Let us estimate the time required to complete the analysis of the entire business area of financial aid. "Award Financial Aid" is by far the largest function performed by the Financial Aid Office; project personnel estimate roughly one-third of all activity is devoted to it. So, multiplying by three the fourteen weeks devoted to analyzing "Award Financial Aid," one arrives at a figure of forty-two working weeks or approximately ten to eleven months of one working year to finish the analyses of the financial aid business area and information system. But, for reasons outlined above, even this estimate is on the high side of the scale and project team members believe one could realistically consider completing the evaluation of this major information system in six to seven months. With additional experience and field testing, this time figure will likely come down even further.

The question remains, however: Is application of the Pitt model cost effective? In determining this, one needs to examine not only the time and cost commitment, but also the quality and usefulness of the products. Regarding the former, project personnel remain deeply concerned about the cost of implementing this model. If this model and methodology are to be widely implemented, the costs of implementation must be brought down. One way to achieve this is by creating a leaner, more streamlined model and methodology. As indicated at various points in this article, project team members believe that further revisions and cuts can be made. Another way to save costs is to conduct more rigorous appraisals. This appraisal can take place at two different levels. One level would be to appraise information systems and to concentrate on the systems with the greatest value, however that value is determined. At a lower level, one could appraise functions within systems and focus the analysis then on functions with the greatest value. Either way, one is "cutting down the field" or at least establishing a priority list of functions and systems on which to focus. The result is a leaner, more cost-effective strategy.

Another way to view the cost effectiveness of this model and application methodology is to compare them to another set of procedures. One valid comparison is with traditional records management methodology. Would it require six to seven months to implement traditional records management activities for the records of the IU Financial Aid Office? There is no question it would require many months of diligent work to interview records creators and managers, to conduct on-site surveys of record series, to identify vital and archival records, and, in general, to design a coherent plan for managing the records of this very large office. However, it is unlikely that it would require six to seven months to complete the work. But what about the actual products created by the two procedures? In this comparison, there is no question that the quality of the products of the analysis using the Pitt model and IU methodology far exceed those produced by traditional records management methodology. The Pitt model and the IU methodology create considerably more detailed and precise information on the functions and transactions of that business unit, on how and why the records are created, on how the system is managing or not managing the records, and ultimately on specific measures for improving the way records are captured, maintained, and preserved by that office. Consequently, in the comparison of the two models on the basis of the time commitment/costs and the quality of the products produced, IU project staff conclude that the Pitt model and the methodology required to implement it compare rather well with standard records management methodology.

Nevertheless, the only responsible answer at this point to the question of whether application of the Pitt model is cost effective, is that it is too early to say, and only considerably more field testing will determine this. But project staff would add that the Pitt model and the IU methodology for implementing it have certainly demonstrated great potential, and the proof of that potential is in the value and usefulness of the products. Project team members have enough confidence in the value of the strategy to justify recommending that IU provide additional funding to continue testing the model, and they would not hesitate to recommend that other institutions seriously consider field testing the model and methodology.[29]

6. What types of skills are required to apply this methodology? Who would be involved in developing the methodology, and what roles would they play?

Experience gained in this project would strongly suggest that a wide variety of skills and experiences, more than any one person is likely to possess, are required to create and effectively implement the methodology. There is no question that the team concept is the best approach, but who should be part of this team? Ideally the team should consist, at the very least, of a

[29] The author has recently applied for and received $10,000 from Indiana University Information Technology Services to continue field testing the IU methodology and the Pitt model.

professional archivist or records manager, an IM administrator, and a systems analyst. The archivist/records manager would bring to the task a working knowledge of appraisal, of records management procedures and techniques, of the value of contextual data, and of the need to create and maintain evidence. A higher-level IM administrator would provide an overall knowledge of all the systems and how they interact, of the individuals who play key roles in managing these systems, and of data management principles and procedures. And finally a systems analyst would bring his/her working knowledge of functional decomposition and of analyzing and evaluating information systems.[30]

Another conclusion derived from experience on the project, a conclusion arrived at by other archivists involved in electronic records management, is that the archival profession needs to add some new skills to its "tool kit" in order to be effective in the world of automated records.[31] Skills that immediately come to mind are a basic knowledge of how automated systems are created and work; a more detailed knowledge of information management methodology; experience with functional analysis methodology and modeling techniques; and knowledge of computer-based information systems, particularly metadata systems, such as data dictionaries and information resource dictionary systems. The goal here is not to have the archivist become a programmer or systems analyst, but rather an archivist who can speak the language of the technologist and who is able to perform some basic tasks related to modeling and describing business functions and data.

7. Will data stewards and managers understand and perceive value in this methodology and model?

First of all, it can be stated without hesitation that IU project staff received very good cooperation from all the automation personnel asked to work on the project. This suggests that data managers and stewards do recognize some value, some payoff for them in the project. One of the benefits most often mentioned was developing disposition schedules for record-creating events. Another major issue which resonated with many within the university was collecting metadata relating to the use of records. Finally, quite a few information managers expressed enthusiasm for a strategy that promised to improve access to more fully documented, inactive electronic records. Overall, project staff sensed among university personnel a growing awareness

[30] Other valuable team members would include someone from legal counsel, a representative from internal audit, and an individual involved in decision support activities.

[31] An extensive body of literature has been generated to identify the skills archivists will need to function effectively in the twenty–first century. In particular see Margaret Hedstrom, "Teaching Archivists About Electronic Records and Automated Techniques: A Needs Assessment," *American Archivist* 56 (Summer 1993): 424–33; and Tom Nesmith, "'Professional Education in the Most Expansive Sense': What Will the Archivist Need to Know in the Twenty-First Century," *Archivaria* 42 (Fall 1996): 89–94.

and recognition that standard back-up procedures were not creating an environment where records could be easily retrieved and were not preserving the kind of records that would be useful in the future.

Did data managers and stewards readily understand the concepts within the model and in the methodology? The answer is both yes and no. On the one hand, much of the methodology was understood very well by systems analysts and information managers, particularly such concepts and techniques as functional decomposition, metadata, and analysis of how transactional data is used, maintained, and distributed within systems. On the other hand, these same individuals were not so familiar or comfortable with such critical concepts as transactions, records as defined by the model, the pivotal role of contextual information, and, most critically, the role and value of evidence in a record system.[32] Based on the experience of this project, there is a very real question as to whether the Indiana University administration is ready at this time to devote valuable resources to projects designed to preserve evidence. Experience would indicate that most of the resources within the IT community continue to be allocated to projects designed to manage data content. To date, strategies for preserving evidence and additional contextual documentation have far less appeal.[33] On the other hand, project staff have witnessed recently a recognition among some high-level administrators that the concept of accountability requires much greater emphasis at IU. Therefore, the IU project team members remain guardedly optimistic.

8. What did project staff learn about the systems they analyzed?

On the positive side, experience with the information systems for the registrar and financial aid information offices revealed that the records are being managed rather well overall. In particular, data managers were very aware of the pertinent laws, regulations, and statements of best practice; and the systems did an excellent job of capturing the necessary transactional records. These offices were also very conscious of quality control, and so the records were generally accurate and consistent. Another major strength of these offices, and likely of most major record creators on the campus, is their emphasis on establishing security measures; consequently, the mechanisms for ensuring that only authorized personnel generate, modify, and delete records were very good. Finally, there were well-established mechanisms for exporting data and for rendering record content as it originally appeared to

[32] In an IT context, a record is defined as a data structure representing an element of a file or a group of related fields. The concept of a business record is not part of the IT methodology.

[33] The author's experiences regarding the priorities of the IT community were reinforced in a recent e-mail discussion with Robert Horton of the Minnesota State Archives. Horton wrote that "evidence per se does not have any immediate resonance with our collaborators, so it can't be used as a starting point or as given. It's not a motivating force or a familiar concept to them so far, it's outside their experience." Quote taken from an e-mail message from Robert Horton to Philip Bantin, 5 June 1998.

the creators. In regard to metadata, except for some notable exceptions, most of the data was present, although it was not always linked to the record or readily available in a way project staff would have liked to see. So, overall, the field tests presented a relatively positive picture of recordkeeping at IU. But project personnel expected this. Of all the systems in a university, those dealing with student records are the most regulated and, consequently, the most controlled and audited. Nonetheless, even with these student systems, serious problems were discovered.

In regard to the functional requirements section, it is fair to say that project staff encountered system shortcomings, some serious and some relatively minor, in all the major areas of the functional requirements. The most serious problems to date are in the functional requirements for maintaining records. In regard to maintenance issues, one of the critical concerns is the lack of procedures and mechanisms for migrating records on a regular basis. Presently, the university creates back-ups of its files, but this procedure focuses primarily on preserving record content and does not properly satisfy many of the key components of a viable migration program. Another major and related maintenance concern is the retention of inviolate records. The information systems that project staff examined do a very good job of preventing unauthorized changes to records, but they do a very poor job of preventing accidental or intentional modifications or deletions by those who have permission to use the system. Another problem typically encountered relates to accountability and, more specifically, to maintaining a responsible system throughout the eight-campus system. The Bloomington campus information systems were very well documented, but there are serious questions about the existence and availability of policies and procedures at the other campuses. Project staff believe that this will be a reoccurring problem with any of the information systems which are not centrally maintained at the Bloomington campus. Finally, project personnel discovered shortcomings in the area of access to the records. Project staff found that although the primary authorized users, usually those in the record-creating units, had excellent training on use of the systems and had ready access to system documentation, those authorized users outside the creating office had far less knowledge of how to access and use the system.

Project staff also discovered several serious shortcomings with regard to the metadata requirements. A major concern is the absolute absence of any metadata relating to use history. Another critical shortcoming is that a good deal of critical metadata is in paper form only, and information on how to access this data is not included in any electronic source, nor is the metadata or a citation to it linked in any way to the record. This problem was found to exist for disposition metadata; for some contextual metadata; for structural documentation that represents meanings, values, and relationships of record;

and for metadata defining terms and conditions for access and use. These categories of metadata are too critical to remain separated from the records they describe. Project staff fear that over time these metadata may accidentally be destroyed or become completely disassociated from the records they document. This is particularly true for those records which are taken "off-line" and migrated to another system. Project personnel also uncovered a problem relating to the "Handle" layer of metadata. It was discovered that although each record has some unique identifier, there was no piece of metadata identifying this set of data as a record which needed to be maintained and preserved according to some well-established requirements.

Finally, project team members suspect that many of the problems the IU project uncovered are not confined to the test cases but will be present in the majority of the information systems on the IU campus. Three system deficiencies are of greatest concern: 1) the lack of mechanisms and procedures for ensuring that metadata remains joined with or linked to the record; 2) no defined mechanisms for migration (as opposed to simple back-ups) of records; and 3) no established procedures for ensuring that records remain inviolate over time. All of these problems are very solvable; but until archivists and records managers can elevate these issues to the top of the IT priority list, these deficiencies will continue to thwart efforts to create viable record-keeping systems.

Future Directions and Next Steps

As mentioned several times, IU project staff are very impressed with and committed to the Pitt model and the method of analysis it recommends. But the model and the methodology to implement it still need more testing before staff members are prepared to present them to university administrators as a strategy ready for implementation.

More specifically, the following issues need to be addressed:

1. In what ways can the methodology for applying the Pitt model be streamlined and made more cost effective?
2. What are the costs based on numerous field tests of applying this methodology and model?
3. How can archivists and records managers more accurately and effectively identify the transactions they will choose to document? At what level of the function/subfunction/transaction continuum will the analysis of record-creating events typically occur?
4. Should archivists be thinking in terms of a priority system of applying these reviews, based on the widely acknowledged value of the information system or of specific functions within systems? And, if so, can

the profession define the systems and functions in such a way that they can be applied to other institutions, particularly those of a similar type, such as academic institutions, health care organizations, utilities, etc.?

5. What are the implications and consequences of the results of other metadata projects for this model? And, relatedly, how do archivists integrate the search for evidence and accountability into other projects.
6. What are the best strategies for integrating this model and methodology into the standard set of procedures undertaken by the institution whenever an information system is created or modified?
7. Can a core set of functional requirements and metadata specifications that need to be used in each and every review be identified?

There are also a whole host of questions related to implementation that need to be addressed:

1. What will be the costs of implementing these recommendations for improvements to the systems?
2. What are the costs of maintaining these recordkeeping systems?
3. How responsive will data stewards and IT administrators be to implementing changes?
4. Who needs to be part of the implementation team?
5. What is the best strategy for joining metadata to the record? Is the strategy of encapsulating the metadata with the record worth the cost of implementation? Put another way, is the risk of losing critical metadata so great that archivists and the IT community cannot afford to implement any other strategy than encapsulation?[34]
6. Where will this metadata be found and how will it be collected?[35]
7. What policies and procedures need to be in place for the model and methodology to work?

[34] David Bearman proposes a strategy designed to create Metadata Encapsulated Objects (MEO's). These MEO's are designed to be self-contained, self-describing, self-documenting objects which logically and physically include all the information needed to reconstruct the conditions which will permit users to read and understand the record as it was created. The best expression of Bearman's strategy for MEO's can be found in his article "Item Level Control and Electronic Recordkeeping," 214–39. An implementation project incorporating the MEO concept into its strategy for creating recordkeeping systems is the Philadelphia Electronic Records project. For more information on the Philadelphia project, see its homepage at <http://www.lis.pitt.edu/~nhprc/perp.html>. For a different version of the MEO approach see Jeff Rothenberg's article "Metadata to Support Data Quality and Longevity" which can be found at <http://www.computer.org/conferen/meta96/rothenberg_paper/ieee.data-quality.html>. Instead of encapsulating all metadata as Rothenberg recommends, another metadata strategy suggests storing the majority of the metadata items in a separate database and linking the record to the metadata via direct links or pointers to the location of the documentation.

[35] Prominent strategies include: 1) extracting the metadata from existing documents—a strategy known as "Data Mining;" 2) manufacturing the metadata and attaching it to the records—a strategy known as "Data Harvesting;" and 3) designing systems which automatically "harvest" pertinent metadata—another type of "Data Harvesting."

8. What are the best strategies for preserving inviolate over time records and metadata which must be used periodically for administrative or research purposes?
9. In what ways can archivists and the IT community automate the identification of record-creating events and the selection of appropriate metadata?

One of the real keys to crafting effective strategies for managing electronic records will be the willingness of archivists and records managers from both big and small programs to step forward and begin testing the strategies and concepts that have already been proposed and that will continue to emerge over the next decade.[36] Until these real life assessments and evaluations are made, the archival and records management professions will find it hard to make any great strides in their quest to address and meet the challenges of managing electronic records.

[36] Margaret Hedstrom makes this same point in *Electronic Records Research and Development: Final Report of the 1996 Conference held at the University of Michigan, Ann Arbor, June 28–29, 1996* (Ann Arbor, 1997), 37.