

**INDIANA UNIVERSITY
ELECTRONIC RECORDS PROJECT – PHASE II
2000-2002**

**FINAL REPORT TO THE
NATIONAL HISTORICAL PUBLICATIONS AND
RECORDS COMMISSION (NHPRC)**

**PROJECT DIRECTOR
PHILIP C. BANTIN**

**PROJECT FUNDED BY A GRANT FROM NHPRC
(2000-036)**

**Permission is granted to quote freely from this report
provided that proper attribution is given.**

**Bloomington, Indiana
2002**

IU ELECTRONIC RECORDS PROJECT – PHASE II

FINAL REPORT TO THE NHPRC

Personnel

Our original plans were not realized. Our initial goal was to have two project archivists. Tom Malefatto, who at the time when the project proposal was written was an Assistant Archivist at the Indiana University Archives, was to be temporarily reassigned to the NHPRC project as the Project Archivist. However, he left the university unexpectedly at the beginning of 2000. Rosemary Pleva, a graduate of the IU SLIS Program, was expected to become Assistant Project Archivist, but once Tom left we moved her into the pool for the lead project archivist. Rosemary began work on the project April 10, 2000, which was over a month later than expected. Once Rosemary began work, I decided to hold off hiring a second archivist until I determined how much work we would be undertaking in the first few months of the project. As I said in an earlier report, and it bears repeating here, in a project like IU's, workload is very dependent on outside partners, such as Internal Audit and IT. Unfortunately for us, a key systems auditor left Internal Audit soon after the project began, and this position was not refilled until early in 2001. This dramatically reduced the number of audits performed in the first year of the project. In addition, we experienced slowdowns and delays in the work on the PeopleSoft project, another major initiative of the project. Taken together this meant less work for us initially, and so we decided not to fill the Assistant Project Archivist position until we actually needed this person.

In the second year of the project, the workload began to increase. At about the same time I learned that Rosemary Pleva Flynn was pregnant, and soon thereafter she informed me that she would only be returning for a short time before she and her husband moved to North Dakota (her husband had accepted a faculty position there). Consequently I requested and received permission to hire another full-time staff member and to retain Rosemary at 50% time to work on policies and other documents from her home in North Dakota. The new full-time position was targeted to be a system analyst/programmer who could work with the OneStart/EDEN project staff on the EDEN workflow engine. Unfortunately, like so many plans related to information management at IU, work on the EDEN workflow engine was put on hold for one to two months as the IT staff worked on other related projects. Essentially this meant that there was no work available for the programmer/analyst, and consequently plans to advertise for this position were put on hold. Although work on the EDEN began in earnest again in February, 2002, it was too late in the project to hire another staff member.

After Rosemary left IU, project responsibilities were assumed by me and by a visiting archivist from Finland, Jaana Kilkki. For the last nine months of the project, I assumed a larger role in all aspects of the project implementation. The percentage of time I

committed to the project increased from 20-25% to 35%. Jaana Kilkki, Director of the Military Archives of Finland, started working on the project in August of 2001. Jaana contributed to fulfilling the objectives of the project in several ways, but her primary contributions were revising and improving the functional requirements, metadata specifications and policy statements. She also gave me excellent feedback on some of our outreach projects.

Finally, we used project funding to hire IU School of Library and Information Science graduates to work on the data survey project and to assist in preparing policies by undertaking literature searches and summarizing other's institutions' electronic record policies and procedures.

Timetable and Products

There can be no disguising the fact that we did not meet all our goals in accordance with our original timetable.

- 1) **Audits of System** (Highest Priority) – We had originally planned on completing twenty-eight audits over the course of the two year project. The goal established by NHPRC staff was to complete at least fifteen audits. We did not meet these goals. During the course of project we undertook twelve audit reviews.
- 2) **PeopleSoft Student Information Module** (Highest Priority) – We had originally planned to create business process models depicting records that were generated, and to submit recommendations for incorporating recordkeeping requirements and metadata specifications into the Student Information System. This goal was not met. We only completed some preliminary interviews in preparation for creating process models, and we never got the opportunity to develop specific recordkeeping requirements for the Student Information System.
- 3) **PeopleSoft Human Resources Information Module** (Highest Priority) - Again, our objectives were to create business process models depicting records that were generated, and to submit recommendations for incorporating recordkeeping requirements and metadata specifications into the Human Resources Information System. We achieved neither of these goals. Our only input on this system was to participate in an E-Docs Subcommittee that reviewed types of forms or documents that would be generated by the system.
- 4) **Document Management Application** (Lower Priority): The goal was to participate in the review of document management software and make recommendations related to recordkeeping requirements and metadata specifications. This goal was met.
- 5) **Generate Electronic Records Policy and Guidelines for Managing Electronic Records** (Highest Priority). This goal was met. During the Project we created four electronic record policies and significantly revised our methodology for evaluating and designing automated systems.
- 6) **Creation of six white papers**. This goal was met. We created seven white papers.
- 7) **Submission of at least two articles about the project**. This goal was met. During the course of the project, we published five articles and gave sixteen presentations on the

project results and findings. My article appearing in *The American Archivist* won the Fellows Posner Award for the year's best article to appear in the journal for that year.

8) Work with Processing Modeling and Information System Review Team (Lower Priority) – The goal here was to complete five process models and reviews of systems. This review team concept never got off the ground, and consequently this objective was not met.

While we did not meet some of our goals, we did undertake some important activities that were not listed in our original plan. These included work on the IU Portal and Workflow System, implementation of a survey of electronic records management activities on the IU Bloomington campus, the creation of a number of policies related to managing digital records, and several revisions of the Functional Requirements for Recordkeeping Systems and of the Recordkeeping Metadata Specifications.

Why is there such a discrepancy between anticipated and actual results? I believe three circumstances contributed to this.

1) Unrealistic goals – Looking back, I believe that the goal we established of more than two audits a month was not realistic. When I made those projections, I was thinking primarily about how many audits my staff could complete, and one audit per person per month did not seem outrageous. What I did not properly assess were the resources the Internal Audit Department could commit to audits of automated systems. As we discovered, IU Internal Audit could not complete audits at that rate, especially in light of unanticipated losses in staff for long periods during the project. It was also unrealistic to assume we could complete all our goals related to the PeopleSoft design and implementation. These were huge projects involving hundreds of people, and our contribution related to recordkeeping was only a small piece. To achieve our goals the timetable for PeopleSoft implementation needed to be completed exactly on schedule, and this just did not occur.

2) Promises not kept or plans that did not materialize – IU personnel managing the PeopleSoft implementation were initially very encouraging regarding our involvement in the project. However, once we started to make plans and request resources, they were not nearly as forthcoming. Part of this reluctance to work with us was undoubtedly a result of schedule changes, but I also believe they were not completely honest with us or did not truly understand our objectives when we had our initial discussions with them about project goals. The Processing Modeling Team never developed as anticipated, and it was nobody's fault. Plans relating to the formation of this team took shape just before we applied for the grant, and once we tried to put the concept into action, it began to fall apart.

3) Unanticipated delays or changes – In many ways, this was the major problem we faced. In almost every initiative undertaken in this project, we were relying on other partners to meet their obligations and maintain the original timetable. We simply could not control our own fate in the same way one can in a processing or scanning project. Throughout the project we were faced with changes in plans that were not anticipated and which had a major negative impact on our project. The most dramatic example is the PeopleSoft implementation. The implementation got way behind schedule, and the emphasis became one of simply completing basic functions and getting the system to

work. With IT managers scrambling just to make the system function, it proved impossible to divert their attention to other issues. Recordkeeping concerns became a real peripheral issue that was relegated to the back burner. In the audit area, loss of staff within Internal Audit resulted in a dramatic reduction in the number of system audits performed. Shortly after the project began the primary auditor for automated systems unexpectedly resigned. The job remained open for almost a year. Even when it was filled, the auditing of automated systems proceeded at a much slower pace than normal as the new person eased into this new job. The other contributing cause was the unanticipated number of audits which were limited to very technical issues. In these cases Archives and Internal Audit staff agreed that participation in the audit by Archives staff was not warranted or useful. Again, it is an example of fact that the project timetable and workflow was to real degree dictated by the decisions made by the project's partners. Regarding audits, I could suggest the addition of audits in areas of greater interest to the Archives, but I could not dictate or force Internal Audit to add them to their schedule.

One of the lessons that I re-learned was that while the Archives can have a records management agenda, it cannot control and dictate enterprise-wide information management priorities. Throughout the project, we had to respond to what was going on around us. We could attempt to shape the records agenda, but at least for the major projects, we could not control it. The best we could do was to insert our perspective and requirements into the design and implementation process. When the PeopleSoft project was delayed and we could not undertake our initiatives, we found another related project, the IU Portal and EDEN Workflow project on which to work. Similarly, when we were not as busy on audits, I completed more work than anticipated in developing electronic policies, and I initiated work on the records management survey. To summarize, in projects like ours that involve the cooperation and commitment of many other partners to make it all work, it is very difficult to maintain a timeline or to meet all the original goals. While I regret that we did not meet all our goals, I am nonetheless satisfied that we took advantage of all the available opportunities to insert the recordkeeping perspective into the dialogue. I am certain that the activities we undertook during this project and the awareness of recordkeeping requirements we have created will bring about very positive results in the near future.

Advisory Committee

The Advisory Committee met twice in Bloomington, on July 21-22, 2000, and on July 12-13, 2001. At the first meeting the major topics of discussion were the partnership with internal audit, the EDEN Workflow project, and the draft policies on various aspects of electronic records management. Terry Radke and James Reinhard of the Audit Department attended the meeting to discuss cooperative audit projects, and Shawn Mitchell of IU Information Technology attended to discuss the EDEN Workflow project. On the second day, the Advisory Committee also reviewed the draft electronic records policies, the functional requirements and metadata statements, and the process modeling projects.

Rather than waiting until the end of the project, the committee requested to meet again next summer. On the morning of July 12th, the project team gave a general overview of the changes made to the project, and discussed with the Committee some of the problems associated with the project due to changes in partners' schedules and expectations. All the Advisory Committee members agreed that for projects like IU's, where a number of partnership activities were being undertaken, there was a need for some flexibility regarding the project's timeline and output expectations. The rest of the morning was spent updating the Committee on the partnership with Internal Audit and on future audits that would be conducted. Jim Reinhard from IU Internal Audit attended the meeting and fielded questions about the involvement of Audit with the Archives staff. The afternoon session began with a presentation by Rosemary Flynn on OneStart and EDEN, which included an explanation of the use cases that she had been helping the EDEN workflow development team write. She was then joined by Kurt Seiffert, a system architect and EDEN Workflow project leader, and Phil McKown, an analyst/programmer, who is also on the workflow team. Kurt provided more detailed information on the EDEN Workflow engine and articulated how he expected electronic recordkeeping functionality to be incorporated into the design. During the session on the second morning, there was further discussion of the workflow engine, electronic recordkeeping, and the data warehouse/decision support environment. There was also discussion concerning the project's proposed list of products, during which the Advisory Committee unanimously recommended that the project team revise the type of white papers project staff should produce.

At each of the advisory committee meetings, we received very good input on all aspects of the project. In addition, advisory committee members provided excellent written comments on the various policy statements. Overall, however, I must confess that I did not make use of the advisory committee as I hoped or planned. My goal was to obtain constant feedback from committee members over the course of the project. One strategy for achieving this was to forward draft documents to the group as they were produced. To a large degree this did not happen. Most of the input was at the meetings, and little occurred at other times in the project. This was largely my fault; I simply did not solicit the input of committee members as diligently as I should have. Another strategy we employed was to set-up an e-groups communication site on the web. At this site we posted documents and solicited comments. This strategy had limited success. Again, I confess that I did not use the tool as effectively as I should. In addition, only some of the committee members used the e-groups and responded consistently. In sum, while the input from the advisory committee did improve the quality of the project products, the committee had the potential to contribute much more to the project had I been more diligent in obtaining their input.

Implementation Projects:

****Audits:**

During the Project we were involved in twelve audit reviews, some large and requiring a commitment of many months, and some relatively small and lasting only a month or two. In some cases, these audits are still ongoing and have not been completed. The audits undertaken by the Project staff included:

EDEN Workflow System – Recordkeeping Requirements – This activity is described in its own section of the report.

Peoplesoft Implementation – Student Information Systems Application – This activity is described in its own section of the report.

Treasurer’s Department: This is a large office of seven managers, a LAN manager and an office manager. The unit’s automated systems include input and output to/from the Financial Information System (FIS – an enterprise-wide financial system), a department LAN and local databases.

Goal: Review operation from a recordkeeping perspective

Timetable: 6 month audit

Findings:

FIS System: Staff are not sure what records are being retained or in some cases how to get access to older records

LAN - General problems: Naming conventions; duplicate files; multiple versions; e-mail; lack of documentation on procedures; no coherent strategy for identifying records to be placed on shared LAN

Imaging: Lack of imaging strategy

Retention: No clear strategy; do not truly trust retention schedules for financial records because they have not been approved at highest levels of the University

Capital Financing and Tax Management: This is a large office consisting of six managers. Automated systems include input and output to/from the Financial Information System (FIS – an enterprise-wide financial system), department LAN and local databases.

Goal: Review operation from a recordkeeping perspective

Timetable: 6 month audit

Findings:

LAN: Same as for Treasurer’s Office; tend not to use LAN because they do not trust it, and experience has shown they cannot find records they want.

Retention: No clear strategy; tend to keep everything

Financial Management Services: Audit of imaging system for processing documentation supporting administrative travel.

Goal: To ensure that imaging system meets recordkeeping requirements and standards for imaging

Timetable: 2 month audit

Findings:

Summary of report: Recommendations: 1) Create more detailed documentation on the imaging process; 2) Incorporate ANSI/AIIM standards for scanning into their process and documentation; 3) Develop a strategy for monitoring compliance with established procedures and standards; 4) Conduct an annual review of imaging procedures and guidelines; 5) Establish procedures for processing illegible originals; 6) Link disposition metadata to the record; and 7) Establish formal and ongoing programs for training staff in the use of scanning hardware and software

Financial Management Services: Review of imaging system for Accounts Receivable Invoices

Goal: To ensure that imaging system meets recordkeeping requirements and standards for imaging

Timetable: 6 month Audit

Findings: Ongoing discussions about the application and design of an imaging workflow, and about the requirements for an imaging program that ensures the reliability, accuracy, security and accessibility of the digital images it creates.

Athletics Department: Audit of Student Recruiting Application System

Goal: Review the recruiting application system in regard to how well it meets recordkeeping requirements

Timetable: 6 month review

Findings: Had initial meeting to discuss design of the system and to discuss concerns of the auditors and archivist; waiting for documentation and more information on design.

Student Loan Administration: Audit of new Student Loan System

Goal: Review the system to determine how well it meets recordkeeping requirements

Timetable: 1 year audit review

Findings: Have had several conversations with them about business processes and recordkeeping requirements. Issues to address: 1) Review existing audit trails to determine how comprehensive they are; 2) Review how transactional data going back to 1950s is managed – what is captured and how accessible? 3) Review what types of metadata is stored in the “History Journal”; 4) Review procedures for scanning documents; and 5) Review their use of electronic signatures.

Registrar: Review of the management of mission critical documents distributed as digital copies only on the Web.

Goal: To develop policy and procedure statement for the management of University’s web sites

Timetable: 6 to 9 month review

Findings: Reviewed issue with Registrar staff and defined main issues: 1) Develop policy and procedural guidelines for web management; 2) Identify mission critical documents and develop appraisal guidelines; 3) Involve legal counsel and data stewards in discussion; and 4) Develop strategy for gaining University approval for policy.

Next steps: Meeting with a larger group of managers in Registrar's Office, Internal Audit, Legal Counsel, and Data Stewards representatives.

Optometry School: Review of VersaSuite software purchased to manage Optometry Department's Clinic Services business processes.

Goal: To review system to ensure it meets basic recordkeeping requirements

Timetable: 6 to 9 months review

Findings: Reviewed business process flow charts and discussed where records are generated; reviewed the basic functionality of the VersaSuite software; concerns and questions about the software included: 1) No apparent way to access record metadata; 2) Could not determine whether the system created audit logs – Can we tell who has modified a record, and how it was modified? 3) Are older versions of records saved, or does the system simply overwrite older data? 4) Does the system include retention and disposition functionality? and 5) Does the system ensure that records are not accidentally deleted?

Facilities Department: Audit of the Maintenance Management System (MMS)

Goal: To review system to ensure it meets basic recordkeeping requirements

Timetable: 6 month review

Findings: We concluded that the MMS meet few of the basic recordkeeping requirements. Specifically, we recommended that the Department contact the vendor about adding the following in future updates: 1) Ability to capture records; 2) Ability to preserve inviolate records for as long as they are needed; 3) Additional recordkeeping metadata to any future updates; 4) Ability to access, display and manage all components of a business record, including the metadata; 5) Develop a strategy for the retention and disposal of records within the system.

Continuing Learning Network at IUPUI: Review of the Network's SignUp System

Goal: To review system to ensure it meets basic recordkeeping requirements

Timetable: 6 month review

Findings: We concluded that the system did not meet some basic recordkeeping requirements. These included that the system: 1) only marginally meets state and federal laws relating to privacy and retention; 2) does not systematically capture and preserve business records; 3) does not capture essential metadata documenting the content and structure of the record and the context of its creation; 4) does not ensure access to all components of a record; 5) does not include an authorized disposition plan; and 6) lacks adequate documentation for record creators or users.

Overall Evaluation of Problems Uncovered by Audits. Re-occurring recordkeeping problems included systems that:

- 1) Do not routinely and systematically capture records
- 2) Do not systematically retain inviolate records for as long as needed or required by law or best practices
- 3) Do not routinely and systematically capture essential record metadata.
- 4) Do not maintain a logical or physical relationship between record content and metadata.

- 5) Do not maintain a logical or physical relationship between records generated by the same or similar business processes.
- 6) Do not include an authorized disposition plan.
- 7) Do not include adequate documentation for record creators and users

****Process Modeling of Student Admissions and Advising Systems**

The goals of this project were to review business processes, assist in identifying processes that might need to change, design recordkeeping requirements for the PeopleSoft application, and help create an imaging application for the areas of Student Admissions and Orientations and Academic Advising. To achieve these goals, interviews were conducted with the Bloomington Undergraduate Admissions Office, Bloomington Undergraduate Orientation Office, School of Business – Undergraduate Admissions and Advising, Allied Health Sciences, IUPUI Honors Programs, IUPUI School of Engineering and Technology, and Undergraduate and Graduate Programs. Questions that were asked included: 1) What are your business activities; 2) What are the outputs or records you generate; 3) Are changes in business processes required? 4) How would an imaging system help you?

Some of the issues or questions that offices asked included: 1) How will these offices interact with PeopleSoft? 2) What is the timeline for implementation? 3) Who is paying for imaging software and hardware? 4) Will there be common software and procedures between Bloomington and IUPUI campuses? 5) Will OCR software be used? 6) Do the Student Records offices need to be involved in this project? 7) What privacy issues exist? 8) What is needed for accreditation purposes? and 8) Should we all be retaining the same records for the same amount of time?

The ultimate use and implementation of the products of this modeling project are in question, however, since funding for implementation of the imaging system and PeopleSoft application for these areas is not in place. Nonetheless, the information on admissions and academic advising records will be very useful to the University Archives staff when they do recordkeeping audits. Furthermore, since imaging is so essential to these units, there can be no doubt that some program will be implemented in the not so distant future. When that time comes, the process models will prove essential to the creation of a scanning program that meets recordkeeping requirements.

****Review of Electronic Document Management Systems**

From June through September, the Project Director and Project Archivist participated in the review of proposals for the purchase of an enterprise-wide document management system. Initially, we focused on promoting the purchase of enterprise-wide records management application (RMA) software, such as TRIM and Foremost. However, IT personnel in charge of review were unfamiliar with these products, and suspected that this RMA software might not have enough functionality to handle all the information

needs of IU. Consequently, it was determined that RMA vendors would not be included in the first round. They readily agreed, however, that records management functionality would be included in the requirements statement of the RFP. The Archives team generated recordkeeping language for the RFP that included requirements for recordkeeping metadata and audit trails, and the functionality to capture records, maintain inviolate records, and create a disposition plan.

We received responses to the RFP from eight vendors: FileNet, IBM, Matrix, Feith, Treev, VanAusdall & Farrar, Bluebird and Artesia. Eventually, we brought in representatives from FileNet, IBM (accompanied by a representative from TRIM), and Matrix to demonstrate their products. We discovered that most of the software products were seriously lacking in record management functionality, and were particularly lacking in such vital life cycle management functions as a disposition management plan and audit trails. Some like Matrix's OnBase simply stated that they do not have this functionality, but would likely develop it in the future. Two others, IBM and TREEV, relied on third party RMA software packages, specifically TRIM and Foremost, to manage records over the life cycle. These were not truly integrated products, but two individual packages linked by a few common fields. Only one product, FileNet's Panagon, truly offered an integrated package that incorporated most of the recordkeeping functionality we were seeking. However, even FileNet representatives recommended that for long-term management of records a RMA like Trim should be used in conjunction with Panagon.

Eventually, no product was selected due to lack of funding and of the will and desire to spend \$100,000 to \$500,000 on enterprise management software. Document and records management applications are still not a high priority at IU. Although, there is growing recognition that data and records are falling through the cracks, this concern is not being translated into a plan or a commitment to solve deficiencies in recordkeeping by establishing a true document/records management environment. Unfortunately this leaves a big hole in my strategy to manage IU electronic records. The IU portal and EDEN Workflow strategy offers an excellent mechanism for capturing records. However, this strategy will have very little impact, in fact it will not get off the ground, if an environment for transferring and managing the records is not in place.

****Policies:**

As indicated earlier, we did more work in this area than anticipated, largely because of delays in PeopleSoft implementation. Actually, working more intensively on policies turned out to be a real plus in moving the records management agenda forward. One of the real deficiencies at IU is the lack of a solid foundation for recordkeeping, beginning with basic policies. Once I began writing the electronic records policy, I recognized how necessary it was to develop a whole suite of related policies. This meant actually stepping back from electronic records management and creating more fundamental documents on records management in general. First, I needed to establish a policy for management of all types of records, and secondly, I needed to focus records management activities in my office by gaining approval by the Board of Trustee for renaming the

Archives as the University Archives and Records Management Office. Beyond this I needed to develop a group of related policies on imaging and e-mail. All these policies are still in draft form, except for one version of my electronic records policy, which was adopted by the Archivists Group of the Committee on Inter-Institutional Cooperation (CIC - Big Ten universities plus University of Chicago) and is displayed on the CIC website. The Board of Trustees resolution and the general records policy are presently being reviewed by the Library Administration. Once they pass on these documents, I will submit them and likely all the other policies to the Committee of Data Stewards for review and approval. Eventually the policies will be submitted to the Committee on Institutional Data, where, if approved, they will become official University policy.

****Methodologies for Evaluating and Designing Systems**

I was never really satisfied with our methodology created after phase I. I was particularly dissatisfied with our procedures for describing business processes and for identifying where records are created, and with our definition of the level at which the analysis of the system occurs, i.e., at the record level, file level, class level, system level.

Throughout phase II, we methodically reexamined our model for analyzing business processes. Consequently, the Archives staff revised the set of activities undertaken in the methodology, modified the vocabulary used, and redefined the products created. During phase II we also re-examined the level at which the review and analysis of information systems would occur. We recognized that if every record or item had to be reviewed at the record or item level to determine if the requirements and specifications were met, the strategy would simply not work. Just as with paper records, we realized that we had to find a way to manage electronic records at the aggregate level. In phase II, we focused on this issue, and I believe we have resolved the problem to some degree, at least theoretically. Emphasizing the necessity of incorporating classification schemes into the system helped immensely in addressing the problem. With a robust classification system in place, we could then begin to manage records as groups – as files or classes of files. The next problem we will need to resolve is determining how this will work in practice, i.e., decide based on experience which functional requirements and metadata specifications require analysis at the record level, and which can be applied at the level of a file or a class of files. We still have not totally resolved this issue, largely because we have not had enough field tests of the methodology to determine what clearly works and does not work. Theoretically, however, I am comfortable with our present approach.

Another change we initiated in phase II was the creation of two methodologies, one for analyzing existing, legacy systems as recordkeeping systems, and a second for designing new recordkeeping systems. In most cases, designing a new system is a two step process involving the description and modeling of business processes, followed by the insertion of recordkeeping requirements and specifications and the results of your business process models into the design of the new system. Analysis of existing systems is normally a more time consuming, more difficult process. It involves not only specifying requirements and metadata specifications and a list of records to be captured. It also

requires an analysis of how the present system is managing the data. In essence, the process involves analysis of “what is” as depicted by models and system documentation with “what should be” as defined by the requirements, specifications and models. Although the methodologies for the design of new systems and the analysis and modification of existing systems share common elements, we thought it made more sense to create two different documents outlining each methodology.

For more specific and detailed information on the evolution of the IU methodology statements, see the white paper on this topic.

****Functional Requirements for Recordkeeping Systems**

In phase II of the project, we determined that reviewing and refining our functional requirements statement was an essential goal. Experience had shown that before we could begin to talk with IT personnel and others about what records management strategies, we needed first to define what types of functionality a recordkeeping system should possess. Consequently, I felt we needed to spend considerable time during phase II of the project creating the best set of requirements possible that outlined how we wanted the system to manage records.

During phase II, we created three versions of the functional requirements, in 2000, 2001, and 2002. In our most recent 2002 version we incorporated radical changes into our set of requirements. The document reflected my sense that we needed a much more detailed and specific set of requirements if we wanted to produce a useful document that would truly inform system designers. This meant greatly expanding the section on record and metadata capture. Because of the importance of classifying the document, we also added in this version a new section on classification schemes. The value of classification in automated systems addresses my continuing concern about the level of analysis, and how one can represent and appraise electronic records at the class or file level. Another new section in the 2002 IU functional requirements statement deals with the creation and capture of system audit trails. Only in the last year or so have I come to truly understand what audit trails represent and how they differ from other metadata. In essence, audit trails document the activities performed on records and their metadata from creation to disposal. As such, audit trails are essential to establishing the authenticity and reliability of records. The 2002 functional requirements document reflects this importance by the addition of eight requirements on the creation, capture and management of audit trails. Another new section in the 2002 version is on preservation strategies. We felt it was important to call attention to this issue by making it a separate section with its own set of requirements. We also wanted to use this section to make important distinctions between preservation strategies and “back-up” procedures. Finally, we added a section at the end of the 2002 version of the IU functional requirements on integrating strategies for managing electronic and non-electronic records.

For more specific and detailed information on the evolution of the IU functional requirements statements, see the white paper on this topic.

****Recordkeeping Metadata Specifications**

In phase II of the project, our objective was to work continuously on refining and improving our recordkeeping metadata specifications. This job was made both easier and harder by the fact that a great many metadata specifications for recordkeeping have appeared in the last five years. During phase II, we produced two versions of metadata specifications – in 2001 and 2002. In our most recent 2002 version, we made some significant changes. The most dramatic change was in the level of analysis and in the level at which metadata specifications are applied. In the 2002 document, we classified metadata into three categories: a) Metadata to be applied at the at record level; 2) Metadata to be applied at the folder or class level; and 3) Metadata that may applied at various levels depending on individual circumstances and needs. Unlike the 2001 version, much more metadata is placed in the third category. The 2001 version made no direct mention of audit trails. In the 2002 version, we have made a conscious effort to identify audit trail metadata and distinguish it from other types of metadata. Finally, in the 2002 version a few new metadata elements were added. They included:

a) Classification Scheme: We added an element that defined the person or post responsible for maintaining the classification scheme and the class and file structure;
b) Relationships to other records: We added metadata elements defining the authoritative record, attachments or appended notes, and aggregation level at which the record is being controlled;
c) Preservation History: We added metadata describing the impact of the preservation strategy on form, content and accessibility.

For more specific and detailed information on the evolution of the IU recordkeeping metadata specifications, see the white paper on this topic.

****OneStart Portal/EDEN Workflow Project (Description of project provided by Project Archivist, Rosemary Pleva Flynn)**

Background

In accordance with Indiana University's Information Technology (IT) Strategic Plan, the University is replacing or re-engineering its enterprise wide systems. This includes the purchase and implementation of PeopleSoft's Human Resource Management System (HRMS) and Student Information System (SIS). The complete project is to be finished by 2004. Once the process began, University IT staff recognized that the replacement or re-engineering of the enterprise wide systems provided an excellent opportunity to integrate all of the enterprise-wide systems at IU. Consequently, an integration plan focusing on a university-wide portal was initiated. According to this plan, access will be provided through a coordinated, unified front end currently called OneStart. Another component entitled EDEN (Enterprise Development Environment) will incorporate the infrastructure that is comprised of components to be shared among applications, including a workflow engine and a global inbox for administrative messages.

In the initial proposal submitted to NHPRC by the IU Archives, project goals as they related to the PeopleSoft project were to establish recordkeeping requirements and

metadata specifications, create business process models, identify records, and incorporate the requirements and specifications into the design of the system. The delays in the PeopleSoft implementation and the emphasis on the portal environment, however, caused project staff to alter their plans. In the first year of the project, construction of the portal became a main focus of the project. The Project Archivist, Rosemary Flynn, was an active participant on the portal development team. In fact, during the first year of the project, she devoted one-half of her time working on the portal team and as a member of the document creation, workflow, and design methodology teams. Her role on these committees was to define recordkeeping requirements, articulate the need for business process modeling, learn how to apply the currently accepted modeling technique for design teams at IU known as Unified Modeling Language (UML), and assist the portal development team create the first release of the portal.

The requirements gathering for the portal began in March 2000 with a Joint Application Design (JAD) session involving IT developers and managers and representatives of several key functional areas including Human Resources, Timekeeping, Student Information, Financial Management, Accounts Payable, Purchasing, Electronic Research Administration, and the Indiana University Information Environment. From this session and additional meetings with these functional areas during the summer of 2000, a vision for the portal emerged. It would provide a unified front end to IU services with single sign-on and authentication and 24x7 availability, role-based customization, usability-tested personalization features, application integration, an adaptive user interface, and a completely user-centered environment. It would also make use of a component-based design (CBD)/Web services methodology and standards approach, which would allow for the development of a shared infrastructure. OneStart can be referred to as a “next generation” portal since it is more than an information portal. Not only is it flexible and responsive to change, it utilizes a distributed model for content creation. The first iteration of OneStart, a pilot for staff terminal services users, went live on May 11, 2001. A second iteration with some additional functionality and a new interface was released on August 23, 2001.

University Archives Involvement

Sometime in the spring of 2000, the director of the PeopleSoft implementation recommended that because of delays in PeopleSoft implementation, Archives project staff would be better served to direct project resources towards the new transaction-processing environment. Her feeling was that records would be created, modified, or deleted by users using forms presented through the portal. These forms would then use the workflow engine in EDEN for routing and approval before updating the PeopleSoft tables. We jumped at this opportunity since this environment would allow us to also capture records created by other applications utilizing the workflow engine. Ultimately, our goal was to build into the EDEN Workflow engine the mechanisms for capturing records along with all relevant metadata as they were created and for directing records to the records management system.

Beginning in May 2000, Rosemary was an active participant on the OneStart/EDEN core development team. Recordkeeping was viewed as both a separate application that would

eventually be integrated like other enterprise applications and as a set of requirements that the portal, workflow engine, and other components needed to incorporate. During the requirements gathering stage, Rosemary worked with smaller teams concerned with document creation and workflow. She articulated our current versions of the functional requirements for recordkeeping systems and metadata specifications and worked with the teams to incorporate these requirements into their design. During meetings with functional areas, she also was able to specify specific recordkeeping needs unique to their systems. During this process Rosemary educated data managers on the importance of electronic records management, and, in turn, managers provided her with basic information on their concerns and issues.

In addition to working with these two teams, Rosemary was also a member of the development methodology team. As part of the NHPRC project, we were stressing the need for business process modeling as a requirement in the design or modification of every new IU system. By participating on this team, Rosemary was able to help shape the development methodology that OneStart, EDEN, and, hopefully, other new applications will use during the analysis and design phases of their development.

From Fall 2000 to Spring 2001, Rosemary continued to work with the portal developers by first writing use cases and then helping turn the use cases into a portal prototype designed to test the feasibility and usability of the requirements. She continued to provide as much technical assistance as possible, including helping with screen design and proofreading sequence diagrams. Once the programmers began coding, Rosemary switched to working again on the workflow engine, work on which had been delayed due to lack of personnel. Rosemary began a new round of interviews with functional areas to determine whether the requirements previously gathered were still valid. She also assisted in developing a new conceptual design for EDEN, explained in more detail below.

Current Conceptual Design

As mentioned previously, OneStart and EDEN were being created using component-based development tools. A component is a specific piece of enterprise functionality that can be reused in future development and integration. Some of the components that are part of EDEN include a workflow engine, recordkeeping, security, and an inbox. A component-based approach has several benefits. First, it creates a repository of reusable business functions that also allows for the replacement of specific functions without affecting the rest of an application. Secondly, it aids rapid application development by assembling existing components and services. Thirdly, CBD can improve the agility, flexibility, and scalability of an application. Finally, as long as components agree upon the protocol to be used (the language and the subject being talked about), an application does not have to reach into another application's database for information. They only interact through their published interfaces.

Workflow is "the automation of a business process, in whole or part, during which documents, information or tasks are passed from one participant [human or machine] to another for action, according to a set of procedural rules." What we are calling a

workflow engine is the software that facilitates that automation. What follows is a very high level overview of the engine and its potential for electronic recordkeeping.

When the IT community and functional areas at IU talk about the e-docs that will be routed by the workflow engine, they are generally referring to electronic forms that are connected to database records. These electronic forms are much like the traditional paper forms they replace. E-docs generally do not include other file types such as scanned images, although the engine will be able to accommodate these. The workflow engine will route e-docs for activities such as completion, approval, or notification.

As far as EDEN is concerned, OneStart is just another application like the Financial Information System (FIS) or the Human Resources Management System (HRMS). To illustrate how the application will work, let us track the routing a prominent document known as the Personal Action Form (PAF). The HRMS will have a component that represents a particular document type such as a PAF. The system will register routing rules with the workflow engine on how this document should route. The workflow engine will route an XML representation of the document from node to node through the business process. The HRMS owns the data and is responsible for its storage and management. However, the workflow engine does tell the HRMS how that should be done in order to maintain the document's route. All communication between the HRMS and the workflow engine is conducted between their respective interfaces.

The process begins when someone initiates a PAF for a new employee named Joe in the HRMS. Whether Joe's PAF needs to be routed to another node for completion or is ready to begin the approval process, the HRMS sends a representation of the document to the workflow engine when it is ready to be routed. The PAF undergoes a discovery process to find the node where it needs to stop next based on the routing rules mentioned above. Nodes will have an associated inbox, which may be for a person, machine, or another process. In this example, the document is sent to a person's inbox by the workflow engine. The inbox is viewed through the portal. At this point, the employee will only have a list of available documents. After the employee selects Joe's PAF from his or her inbox, the request is sent to the workflow engine. The workflow engine will call on a HRMS component for a representation of the document. The component will request the workflow engine to check the status of the PAF (the document's state and what actions can be taken) and will then provide a representation of Joe's PAF that is displayed through a OneStart channel. After whatever action is taken, the discovery process begins again until the PAF reaches its final node.

In this conceptual design, recordkeeping functionality would be added by including a recordkeeping node. To capture the document when it becomes a record or evidence of a transaction, routing rules would direct the document to that node. The attached inbox would have a conduit that passes the document into a recordkeeping system or repository to be managed. Each document in the recordkeeping system would include the metadata attached at its creation, in addition to all metadata gathered whenever it is routed through the engine. Since the record has been captured as part of its associated process, we can be

assured that the metadata also contains all of the appropriate contextual information that may be missing if captured at a later date.

So, where does IU's portal efforts fit in with those at other colleges and universities? Several schools have developed or are developing portals, but their numbers are still few. Most of the portals are simply information portals without the middleware such as an integrated workflow engine. Some of the universities with more developed portals are just now talking about adding this middleware layer. We currently do not know of any college or university that is trying to incorporate recordkeeping into their portal and shared infrastructure in the way we are proposing at IU.

****Data Management Survey:**

A major obstacle we face in managing IU's electronic records is the lack of knowledge regarding the types of electronic data and information being created, and how it is being managed. Consequently, I requested and was given permission to use project funding to survey University offices to determine what data is being created, used, and managed within these units. There was much support among data stewards and managers for conducting this survey, since it would provide invaluable and vital information for planning future data, information and records management projects. To conduct the interviews, I hired several interns from IU's School of Library and Information Science (SLIS) and, as a cost share, I asked two full-time Archives staff members to participate.

We created two survey forms. The first form was mailed out to administrative assistants in the units, and requested that they identify the categories of data or records that their office routinely creates. Once we received the response, we followed this survey up with a second more detailed form that elicited specific information on the characteristics of data, who managed it, and how it was managed. This information was collected by means of one or more face-to-face interviews with personnel in the unit. Overall, we sent the initial survey form to ninety-eight University offices and received responses from twenty-one offices. We followed-up in all cases with an interview and completion of the second survey. Once the second survey form was completed, I met with fifteen of these offices to review the responses and discuss strategies for managing their automated systems.

Here is what we found from the surveys and personal visits: 1) Units were retaining electronic records for too long, particularly financial, personnel and student records; 2) They were converting many electronic records to paper or were creating shadow electronic systems or databases in which they inputted data about certain transactions, because they feared that they could not get this data out of the central systems. Sometimes this fear was based on real experiences in which they requested data from a central system and were told it was "archived" off-line and could not be retrieved for one to two weeks. They justifiably felt this response was not good enough, so they determined they must become more self reliant. The motivation for other individuals was not based on real experiences, but rather on a presumption that it would be difficult to

retrieve data, or on a lack of understanding about who was keeping the data and for how long; 3) This brings me to a related point. Personnel in most units did not understand how information and data moved or flowed throughout the University, and consequently did not know who retained original records, and how long originals and copies needed to be kept. This was especially prevalent for financial, student and personnel records; 4) Local records managers did not have a good grasp of how the systems they were responsible for worked and how they should be managed, and they were unaware, on the whole, of the challenges associated with the long-term management of digital objects. 5) In general, I found that personnel in units wanted to do the right thing, but they did not have the information or skills to meet the challenges. They tended to save too many files, to convert electronic records to paper documents, and to duplicate data files to ensure that they would have access to data and could produce the reports they need. The most important needs are for retention schedules, for education in managing digital objects, and for instilling in managers a better sense of how information flows throughout the University.

Outreach - Web, Papers, Presentations

We have been very active in publicizing the results of the project. Listed below are the outreach activities we have undertaken.

Web Page: We have mounted all our important products on the Electronic Records Project website at <http://www.indiana.edu/~libarch/ER/>

PRESENTATIONS, 2000-2002

By Philip Bantin

“Implementing an Electronic Records Management Program” at Medical Library Association Meeting in Dallas, May 20, 2002

“Creating Collaborative Policies: Taking the First Steps in Developing an Electronic Records Program” at Midwest Archives Conference Meeting, Minneapolis, May 4, 2002

Workshop on “Implementing an Electronic Records Program” for CIC University Archivists, Minneapolis, May 1, 2002

Workshop on “Implementing an Electronic Records Program” at ARMA, Utah Salt Lake Chapter, April 24, 2002

“Lessons Learned from the IU Electronic Records Project” for a meeting of the Advisory Committee of the University of California, San Diego Supercomputer Center NHPRC funded project “Methodologies for Preservation and Access of Software-Dependent Electronic Records,” March 11, 2002

“Strategies for Implementing an Electronic Records Management Strategy: Lessons Learned from the Indiana University Electronic Records Project” at a Conference entitled “E-Records Conference 2001” sponsored by Tarian Software and the Houston ARMA chapter, Houston, November 5, 2001

"Protecting Organizational Information: Developing Partnerships for Managing University Information Systems," EDUCAUSE, October 29, 2001

“Strategies for Managing Electronic Records: Lessons Learned from the Indiana University Electronic Records Project” at a Conference entitled “Preservation and Access for Electronic College and University Records” (ECURE) held in Phoenix, October 12, 2001

“Document Archiving and Electronic Filing” at a Conference of the Treasury Management Association of Indiana in Indianapolis, September 20, 2001

"Lessons Learned from the Indiana University Electronic Records Project," Society of American Archivists (SAA), September 1, 2001

“Presenter and Panel Discussion Member for session “”Point-Counterpoint: Do Institutional Archivists and Manuscripts Curators Belong to the Same Profession Anymore?” at Midwest Archives Conference meeting, Cleveland, October 20, 2000

"Strategies for Managing Electronic Records: Lessons Learned from the Indiana University Electronic Records Project," Conference entitled “Preservation and Access for Electronic College and University Records” (ECURE) held in Phoenix, October 5-6, 2000

“Electronic Records Management: Training Needs and Issues,” National Forum for Archival Continuing Education (NFACE) Conference, April 27-28, 2000

“Electronic Records Management Strategies” at a meeting of the Association for Records Managers and Administrators (ARMA), Indianapolis Chapter, March 14, 2000

Presentations by Rosemary Pleva Flynn

"OneStart/EDEN: A Description of IU's Transaction Processing/Recordkeeping Environment," Society of American Archivists (SAA), September 1, 2001.

"Protecting Organizational Information: Developing Partnerships for Managing University Information Systems," EDUCAUSE, October 29, 2001

PUBLICATIONS, 2000-2002

By Philip Bantin

Records Management in a Digital World, a Research Bulletin for the EDUCAUSE Center for Applied Research, 2002.

“Electronic Records Management - A Review of the Work of a Decade and a Reflection on Future Directions” The Encyclopedia for Library and Information Science, Vol. 71, Supplement 34, pp. 47-81.

“The Indiana University Electronic Records Project: Lessons Learned,” The Information Management Journal, Vol. 35, No. 1 (January 2001): 16-24.

“Strategies for Managing Electronic Records: Lessons Learned from the Indiana University Electronic Records Project,” Annotation, Vol. 28, No.4 (December 2000): 18-19.

“The Indiana University Electronic Records Management Strategy – Revisited,” The American Archivist, Vol. 62, No. 1 (Spring 1999): 153-163. Actually submitted in 2000. Winner of the Society of American Archivists’ Fellows’ Ernst Posner Award for the Outstanding Article published in Volume 62 of The American Archivist.