



Turnaround document: Presentations

Educating archivists and their constituencies
State Archives Department
Minnesota Historical Society
16-17 March 2001

Table of contents

1. Project overview
2. What do we mean by XML
3. What do we mean by metadata
4. Indiana University
5. Delaware Public Archives
6. Ohio Historical Society
7. Kentucky Department for Libraries and Archives
8. Smithsonian Institution Archives
9. Minnesota Historical Society

NB: This document is not meant to be a complete transcription or a polished report of the meeting in St. Paul. The goal was to record the sense of the participants' contribution in order to facilitate a common understanding of the project and a consensus on its work plan. If more finished texts were available, they were pasted into the document.

Charlie Arp (CA)
Phil Bantin (PB)
John Churchman (JC)
Rosemary Pleva Flynn (RF)
Edie Hedlin (EH)
Bob Horton (BH)
Richard Marciano (RM)
Glen McAninch (GM)
Ed McNeeley (EM)
Charles Robb (CR)
Shawn Rounds (SR)
Tim Slavin (TS)

Project overview

(Bob Horton)

To start us off today, I'm going to talk about two things:

- what we plan to do in this project
- what we plan to do in this meeting

In the process, I'm going to talk about some of our experiences in MN's electronic records program, touching on the subjects that we asked you to discuss in your presentations.

The ER program here started pretty much from scratch about 4 years ago:

- we had to meet people
- we had to find out which opportunities were available
- we had to educate ourselves – we didn't know much about IT
- we had to educate others – they didn't know much about archives

There were some good models locally for us to study. The GIS community was enormously impressive – very active, intelligent and cooperative group of people. What characterized their work – their success? They had an incentive – they wanted to share their products, their data sets. But they nurtured that incentive and emphasized cooperation. They formed a community, with ties established by continuous education – workshops, meetings, resources – and ties reinforced by practical tools – applications, metadata, clearinghouses. The morals we drew from our own experiences in developing an electronic records program echoed this example.

First – community. Clearly, our best results are a product of some ongoing mutually beneficial relationships we have developed. we could not advocate, establish or produce anything without the support of agencies and people in IT and management.

We built the foundation for that support through both formal and informal contacts. We diligently attended and participated in the Information Policy Council (the IPC is an organization for government agency CIOs). we helped establish and nurture two standing organizations, the Data Issues Group for Information Technology (DIG IT – we didn't help name it), which brings together systems and data administrators across government. And we helped establish and nurture the Government Records and Information Network (GRIN – we didn't help name that one either. Personally, I think GRIM or perhaps GRIME would be more appropriate). In any case, GRIN brings together records managers in government.

We worked with any number of ad hoc groups and task forces: countless metadata efforts (with the most recent being the most successful – Shawn will discuss that); information policy council study groups, one of which helped sponsor our Trustworthy Information Systems project; the Electronic Real Estate Recording task force; a data warehouse study group; a variety of focus groups on a number of subjects ... Sometimes these are all just the same 6 people in a different room, with a different agenda – even so, these are the same 6 reliable and cooperative people

forming the nucleus of a community that lends us invaluable support. In all these efforts, what we did was make contacts, get our message across and form partnerships that led to successful projects and useful products.

I mentioned “incentive” earlier. The GIS community had an incentive – the members all wanted to share data. We have noted a variety of incentives to work together. One that is a most attractive carrot is education. Everyone always needs to know more about IT. This is a very strong catalyst to get together. To foster education, we sponsored all sorts of things – most notably in this context, the XML presentation by Richard marciano and Chaitan Baru last year.

Education is a necessity for us – we have to learn. Education is mutually beneficial – our partners really appreciate our efforts to share knowledge. Education is an element to the program we can manage and sustain. And education can always carry an extra message – even if it’s only “these archivists know how to cater a meeting.” Most often, though, education is an effective vehicle to make our concerns known. People learn about the archives and about archival issues as part of the educational package.

And I mentioned practical tools. We don’t have anything like ArcView. But as we explored the possibilities, we kept coming back to some basic concerns – whatever people wanted to with their data and records, two things would help them immensely: metadata and XML. We heard about these over and over again.

We put all these ideas and factors together in the proposal for this grant:

Community. To meet the challenges posed by IT, archivists needed partners. They needed the support of a wider community of customers and constituents to get the resources, the help and the information essential to an electronic records program.

Education. To understand those challenges and develop responses, everybody in the community (and especially archivists) needed to learn more. Education would bring them together and help them to work together.

And, finally, to be effective and practical, tools. The educational process could focus on two tools that were essential to any records, data, information or knowledge management program – XML and metadata. And let me stress that. These were common denominators. Whatever anyone wanted to do, these would be practical, useful tools.

That adds up to the proposal we made to the NHPRC:

This project will address a critical responsibility that archives have discovered in their work with electronic records: the persistent need to educate a variety of constituencies about the principles, products and resources necessary to implement archival considerations in the application of information technology to governmental functions. It will focus on two issues: XML and metadata. By focusing on these, the project will accomplish several goals. It will:

- raise the level of knowledge and understanding of essential electronic records skills and tools among archivists
- help archivists to reach the electronic records creators who are their key constituencies

- provide the means to form with those constituencies the communities of learning that will support and sustain collaboration
- raise the profile of archivists in their organizations and promote their involvement in the design and analysis of record keeping systems

And everyone here has agreed to help. The work plan in your packets details the steps we want to follow over the next two years. Here are the products we want at the end:

The project is designed to create a number of educational resources. The most important are two workshops, one each on XML and metadata. The workshops will be supplemented with more detailed information, readily available for further research and application, including basic introductions to XML and metadata. The full curricula for the workshops will be available for use and adaptation. In addition, as the available knowledge about metadata and XML will continue to expand, an annotated list of resources will allow for independent research and development. Finally, case studies of implementation within the partners' organizations will demonstrate the practicality of the efforts and their potential for adaptation. The Minnesota Historical Society will maintain a project web site that will provide online access to all these products, as well as ongoing reports on the project.

So how do we start? The basic purpose of this meeting is information gathering and sharing. Because the educational resources we create have to be practical and portable, we thought it best to start with a careful analysis of the status quo. We defined the status quo narrowly: in the email request you rec'd, we asked you to focus on these issues:

1. XML - How have you learned about XML? What XML-related projects are you working on?
2. Metadata - How would you define metadata? What standards are you using or learning about? What projects are you applying them to?
3. Education and communication - Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in anyway to facilitate education? What's your preferred way to learn?
4. Constituencies - Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

We want to discuss these issues in your presentations and then we want to move on to some basic analysis. We want to identify especially what we have in common, because that will be the foundation on which we can build. At the end of the meeting, we should have a common understanding of where we're starting, a consensus on our priorities and a clear view of where we're going – at least in terms of the next steps to take as we prepare for our next meeting in the fall.

To do this, we're going to use (more or less) a process known as JAD – joint application development or design. JAD was first developed by IBM, as a means to facilitate the design of computer systems. “Joint” comes in because the process brought together the IT and the business community. The idea was to avoid the “notes thrown over the wall” approach to applications design that characterized relations between these communities. They didn't communicate well. They wasted a lot of time. They didn't establish a consensus on goals or work plans. And many, many times, they didn't succeed. The IT people were going off in a corner, creating an

application, handing it over and then hearing lots and lots of complaints from the people who had to make it work.

Enter JAD. JAD is supposed to achieve a number of things:

- it facilitates discussion
- it structures discussion
- it documents decisions
- it builds consensus
- and, most important, it fosters success

We encountered JAD in our modeling work and we were impressed. It made modeling possible – or tolerable. We adopted it (and adapted it) for our work with groups because we like consensus and we like success. It was an important feature especially in our tobacco records projects. The main elements we’re going to emphasize here are structure, documentation and consensus.

I’m going to facilitate the discussion, which in practice will mean that I keep it on track, keep it structured. That will involve a number of things – some nagging, some questioning, some encouraging and reminding – I’m going to keep us pointed towards the goals of the meeting.

One tool that I’ll use – a parking lot. We have here a group of smart, articulate and independent-minded people. This is a recipe for chaos. It’s inevitable that a lot of important ideas and issues will come up. They may not all be pertinent to our goals. So when something important but not directly relevant comes up, I’m going to record it in the parking lot as a “side issue” and ask us to move on. We can come back to the list at the end of the meeting and analyze these – we’ll see if, on second thought, they are directly pertinent and, if so, what we need to do about them.

The real work in all this will be done by Jennifer, who is keeping a running account of the discussions. She’s working from a template – the structure for the meeting, those questions just noted. She’ll take down an abbreviated version of your presentations, fitting comments into the template. Then, as we move into a general discussion and start making decisions and talking about our next steps, she’ll record those as well.

Shawn will act as the “coach” during all this, supporting what Jennifer and I do, keeping us on track, pointing at her watch, filling in whenever and wherever necessary.

If you look at the agenda, you’ll see how this works in practice. We’re starting with three introductory presentations- that’s to give us some common vocabulary and raise what we consider to be the pertinent issues. During the presentations, the project staff will look to categorize and evaluate the issues that you raise – I’ll try to point the discussions towards the four issues we talked about earlier, while Jennifer and Shawn sort and record them. At various points, the three of us will analyze what we’re hearing and come back to you with our questions and concerns. We want to be able to react what you’re saying, to verify what we’re hearing and then to try and see how it can be worked into the framework of the project (or, conversely, to see if we have to adjust the framework).

At the end of the meeting, the project staff will take Jennifer's notes and create a "turnaround document." We'll add the materials you have supplied for your presentations, put an introduction and conclusion on the file, clarify what we can and then turn it around, sending the file to you, as quickly as possible, ideally by the of next week.

You'll have a week to read and suggest any clarifications, revisions and improvements. We'll incorporate those and the end result will be a pretty detailed record of what happened, achieved all while everything is still fresh in our minds. More importantly, it will be a record of all that we agreed upon. That will be the blueprint for the work we do over the next few months. We'll return to it at our next meeting in Indianapolis at MAC.

Last but not least, we'll end the meeting with a brief evaluation. Since we plan to meet again and to use this technique, we'd appreciate getting some feedback from you about how things went. So we'll ask two questions: what did we do well? and what could we do better?

When we close up shop tomorrow afternoon, we should have accomplished several things:

- we will have set priorities – what to do, to learn, to produce as our very next steps in the project.
- we will have set means of communication – how we're going to use email, the web, reports and so on.
- we will set expectations and responsibilities – we should all understand who's going to do what as we prepare for the next meeting at MAC in Indianapolis.

Finally, we'll use the project web site for documentation and communication – the URL is <http://www.mnhs.org/preserve/records/edarchivists.html>. Everything we do will end up posted there.

What do we mean by XML?

(Richard Marciano)

Target audiences, constituencies packaging information for?

How get away from these style of presentations? To something more pedagogical, interesting,
better ways to present the information

PB: in a context archivists better understand

Core XML technologies

Future of XML (end of April, Society of California Archivists (SCA))

Run through sample agendas and how evolve from general to more specific audiences

SCA 2001 (draft)

Workshop, intro to XML, themes more useful to archivists

Sessions build on each other "five progressively more complex presentations":

Core: Explain what XML is

Creating Data in XML

Retrieving Data in XML

New directions for XML: standards, EAD and XML, NARA, knowledge-based, topic
maps, trends, what no longer useful

[directions more relevant for archivists]

Discussion: Wrap-up session

Information Integration Using XML (presentation in Minnesota, March 2000)

See agenda on State Archives web site

XML standard, supplemental standards and technologies

Difficult to emerge with building blocks

Questions heard: How to use back at the office, what tools do I need, what education do I need
to use XML, is it worth the investment of time/sweat/blood and tears?

NHPRC XML Tutorial

Daylong workshop, Supercomputing 2000 Conference (Texas, October 2000)

AM session:

XML core

Querying and transforming XML (transformation languages, customized views of underlying
XML data)

Demonstration Session (address shortfalls of MN workshop-tools, applications people can
use) (standards, tool types, products) to address earlier void, examples of what can do
with XML-step to putting XML into action

PM session (application areas)

Data and information modeling with XML (hooks XML offer to map data models or business
processes into XML)

XML vector graphics and GIS applications

Wireless XML applications (WAP) [delivery of XML content]

Focused community=tools with more functionality, versus tools for broad audiences

Not just talk about standards, how people use standards to build useful applications

Application-building

Feedback: interesting-sense application domains to which standards and tools be applied, still want to be able use tomorrow/ confusion how go about this

NHPRC Project (San Diego, March 2001)

“Methodologies for Preservation and Access of Software-dependent Electronic Records”/

Archivists workbench

Apply XML standards and tools in integrated workbench, incorporates alleged functionality archivists may be interested in

PB and RH on advisory board

Agenda: (presentation groups/themes)

Accessioning workbench processes

Accessioning issues for complex documents and the vision beyond

Archival repository issues

Reference workbench issues

Context creation, provenance-metadata standards, knowledge metadata-based standards

Workbench tools, archival tools

IDEAS

Tailoring XML specific constituencies-presentation and mapping issues

Ferret out content relevant to particular audience

PB: his community XML and data modeling, harder approach

Not just recycling presentations

Core XML technologies converging and stabilizing-base to build on

Convey clearer/concise way core technologies, verticalness of community, how core techs connect with pressing issues in field

Multi-modal presentation-wireless/low bandwidth

Presentation of Core XML standards and update on technology (based on October 2000)

GM: Touch on EAD?

PB: Presentations to constituencies would begin with the core

Assume standards change and evolve, but still rely on standards-map old to new

GM: Metadata evolution: MARC, EAD, Dublin Core

PB: Add record keeping metadata

(How will XML capture the data)

What XML is-intro

Archivists-SGML, computer science-XML-two not often relate well together

History XML and SGML/HTML (comparison)

Further reduction of XML standards-normalizing/simplifying, and reversal to complexity

DTDs-specify classes of XML documents, ex. EAD

definitions, hierarchies

Parsers

XML Schema

EM: Mark-up language or processing language?

CR: What communities apply simpler versions?

See OASIS, XML Cover pages-various industries
E-Commerce initiatives

PB: Which web page is this on?

San Diego, buried under publications, RM will resurface for us [provide links in turnaround document to aforementioned presentations, if available]

PB: Some of the examples may be useful for constituencies.

Skip syntax/actual markup

PB: How much detail do we present in workshops?

GM: Focus simple documents/examples (e-mail)/records need to be captured.

JC: Version of XML be understandable 15-20 years from now?

RM: Standards will change, map old to new.

JC: Backward compatible?

RM: Who involved, how communities take techs, how open standards remain, compliance issues

BH: What flows out of grant is voice for archivists.

What do we mean by metadata?

(Shawn Rounds)

I would like to take some time this morning to talk about metadata in terms of both our past projects and a current effort we've got going to develop a recordkeeping metadata standard. The work that the Minnesota State Archives is doing now has its foundation in our recent NHPRC grant project which established our electronic records program. The centerpiece of that project was the *Trustworthy Information Systems Handbook* which offers guidance to government agencies on how to produce and maintain reliable and authentic records.

The development of the Handbook was neither fast nor easy. We had an idea of what the problems were with electronic record keeping in state government, but we had to find out what the people in government thought – we needed to know what challenges they felt they were facing and what they thought they needed to meet those successfully.

To help accomplish this, we held focus groups and workshops, and hosted informational sessions. Some of these focused on a specific topic like metadata, data warehousing, or XML, and included a wide range of participants; others targeted specific groups like data warehouse practitioners. Sometimes we threw the invitation into the wind and let whoever was interested attend; other times, we carefully chose who to involve based upon their experience. We were able to do this – to have access to large numbers of people as well as knowing specific people – in large part because of our heavy involvement in various state government special interest groups like the Information Policy Council, DIG-IT, and MN GRIN.

As we held these sessions, we not only heard about people's needs, we also helped them gain a common vocabulary with which to talk about their problems with us and with others. What we learned very early on is that "metadata" means different things to different people.

To an IT person like a data administrator, metadata is information like field size, allowable values, and file structure. A data warehouse practitioner would list additional items like information about the data's source system, and information about whether the data was cleansed and/or transformed when it was loaded into the warehouse. In both cases, information which supports the operation of the system.

A business analyst generally takes a broader view and includes such things as business rules and data models – information that goes beyond operation to understanding the "why's" and "how's." For instance, why does one department share information with another? Why do we collect information on individuals and how is it represented in the system?

A records manager, faced with an information system or data warehouse, probably wants metadata at least on when a record is created, who creates it, what record series it fits into, how long it needs to be kept, and what is supposed to happen to it after the retention period. A librarian wanting to catalog that record for public access would probably care little about retention information, but would add metadata like "title," "keywords," and "description."

All of these flavors of metadata have their uses and values. But, although each is different, the important point is that all serve the same basic purpose: to allow management of, and access to, information, data, records – whatever you'd like to call it.

It was the practical utility of metadata that kept people interested in talking to us. Our partners in government are trying to run a business. They want practical information and tools that they can take back to their agencies and apply to their work. We learned early on that no one was really interested in theory or in helping us put records on a shelf in an archive somewhere. This message has been reinforced for us time and again from the lowest levels in agencies up to the CIOs – they need to send a tangible benefit, often in the form of positive return-on-investment.

As I mentioned, the TIS Handbook helped us not only to identify needs, but also to meet people through our various sessions, through our testing, and through our promotional activities. These people have been a tremendous resource for all of our work, offering us expert advice, serving as guinea pigs for our products, and working with us as partners. We judge our success largely in terms of our relationships with them. It wouldn't have been possible to continue these relationships if we weren't consistently turning around and giving them practical tools and assistance. For instance, we published the first version of the TIS Handbook last year. Since then we've supplemented it with a legal risk analysis tool and a records management risk guide. We offer several online resources on topics like XML, metadata, search engines, and data warehouses. We also serve a clearinghouse function for state government data warehouse practitioners, and we help administer and host groups like DIG-IT and MN GRIN.

Throughout all of this, we've been using the Web as our primary means of communication. By posting items on our web site, we can reach a wide, dispersed audience very quickly and easily, and we can offer updated, current information almost as fast as we get it. We usually offer people the option of an online, HTML version, as well as PDF file. This self-serve method has proven quite popular. Since its publication just over a year ago, the TIS Handbook has consistently been in the list of top ten downloads for the Society, averaging over one hundred downloads a month. A recent study report, that I'll talk about more in a minute, actually cracked the top five in January, with over two hundred downloads.

I guess I've taken the long way around to talking about what we're doing now. One of our current projects is an effort to develop a record keeping metadata standard for Minnesota government entities. As you've heard, we've been forming relationships with various partners and we've all gotten pretty good at speaking in the same terms, or at least at recognizing our differences. The impetus for this standards project came from our partners and from people in DIG-IT and MN GRIN who identified that they had a problem and went looking for a solution.

Everything got underway late last year when DIG-IT, MN GRIN, and the State Archives recommended to the Information Policy Council that a study group be formed to look at the problem of record keeping metadata and to see if perhaps the record keeping metadata standard developed by the National Archives of Australia might be suitable with modifications for use here in Minnesota. The IPC agreed, with the understanding that this would be a short-term group, coordinated by the State Archives, and that participants would be nominated by agency CIOs. From our past work, we had an idea of who we wanted involved, and we lobbied to get

those people. We ended up with a good mix of skills and experience – the roster included data administrators, data warehouse practitioners, business analysts, records managers (including the state records coordinator), a legal expert, a Dublin Core web-tagging expert (who was also a librarian), and a GIS expert.

The Australian standard was chosen for study because it's probably the most complete government record keeping metadata standard in use. It was first released in 1999 and since then it's been getting a lot of international attention. Its Australian companion piece, a record keeping management standard, is in the final stages of becoming an ISO standard. An accompanying metadata standard is under discussion as well, and it's very likely that the Australian standard will be a foundation piece in the creation of that.

Another item in its favor for us is that it meshes well with the Dublin Core element set and our state GIS metadata standard. Thanks to a project administered by the Minnesota Department of Natural Resources, the Dublin Core has become widely adopted by agencies for tagging web-related resources. Similarly, our state's Land Management Information Center has developed a GIS metadata standard, based on the federal standard, and has been very actively and successfully promoting its use among government entities in Minnesota at all levels.

The record keeping metadata study group met several times from September through December, looking at the Australian standard in detail. The first order of business was to establish why we were undertaking the effort – What would a record keeping metadata standard do for government agencies? Was there really a need for one? To answer these questions, the group developed a business case for record keeping metadata and metadata in general.

We came up with a list of several items relating not only to records management, but also to the daily operational needs of agencies. For instance, record keeping metadata can assist in locating and accessing records for business purposes, such as compiling a report, as well as to meet statutory requirements like properly fulfilling a public access request. Records management can be improved with metadata because it can hold information about such things as retention periods, storage and preservation requirements if the records are moved out of the active system for any length of time, and proper access with respect to our state's data practices law. Such metadata can also facilitate data sharing, a hot topic right now in Minnesota, given a strong push by our governor for electronic government services for citizens. Whether within an agency, or between agencies, data sharing is only possible with consistent metadata that describes the records, their appropriate uses, and who can have access to them.

Once the business case was established, the study committee moved through the Australian standard element by element, holding each one up to the business case to see whether the element supported it, if so, whether it was usable in its current form, and noting any changes that would have to be made.

At the end of its work, the study committee decided that, in fact, a record keeping metadata standard is necessary and that the Australian standard would provide a good model for a similar one here in Minnesota. The group compiled its findings into a final report that was presented to the Information Policy Council in January. The group recommended that the IPC form a

development committee to actually undertake the task of producing a Minnesota-specific version of the Australian standard and that local governments be involved as well as state agencies.

It was not a given that the IPC would support such an effort, for the reason I mentioned earlier – agency CIOs want practical products with a good return on investment when implemented. A couple things were in our favor, though. The membership of the committee had been picked by the CIOs themselves and represented a wide variety of agencies – so support was broad, at least at the staff level. Secondly, the completed TIS Handbook, another product recommended by the State Archives, had earlier received formal approval by the IPC and was, in fact, receiving some good national publicity which reflected very well on the council. With the completion of a formal charge listing potential members (again, nominated by CIOs), a summary work plan, and a list of issues to be considered, the IPC approved the formation of the development committee.

The committee held its first meeting a few weeks ago and over the next year will work on the standard in consultation with the state's Office of Technology, which has the responsibility for maintaining and implementing standards and guidelines. Meetings are open to anyone who is interested because we want to get as much input as possible from as many perspectives as possible. The State Archives is again coordinating the work and, as part of that, is maintaining a committee web site that offers meeting information, agendas, minutes, work products, and links to related items. There's also an online comment form to facilitate communication. Of course, we'll be using e-mail lists for the IPC, DIG-IT, and MN GRIN to help keep people informed.

One committee member, the data administrator for the Department of Transportation has already embraced this standard wholeheartedly and is working on a practical implementation architecture that will allow the standard to work between disparate systems both within Mn/DOT and with other agencies. He envisions a centrally-coordinated state hub that would use non-proprietary software and tools as well as XML to allow communication and data sharing between systems. In this loop, XML would enable this metadata hub which, in turn, would enable data sharing through XML. That's a very simplified summary, but you get the basic idea. What's perhaps more important, he's have success in actually getting people work with him on this project, and not just from within his own agency – people are becoming interested in using the standard in their daily business. Now all the development committee has to do is catch up to him with the standard itself – perhaps easier said than done, but certainly on the way to realization.

Indiana University

1. How have you learned about XML? What XML-related projects are you working on?

Responsibility for XML will fall on Gerry Bernbom. More knowledgeable about IT, will learn it, will be able to understand IT specific terminology

Not a lot of knowledge-need more training. Recognition that XML has a potential, but lots of fundamental work yet to be done.

Expect XML message to be well-received

IU has HPSS – mass storage facility from IBM. No document management or preservation strategies in place. Good opportunity to employ XML. Some awareness of SDSC projects among staff.

2. How would you define metadata? What standards are you using or learning about? What projects are you applying them to?

Generally very good understanding of metadata. Experience and expertise gained over last 5 years. Everything focused on record keeping, what we do as archivists.

Archives can easily go out and teach. Developing metadata requirements for record keeping. Still refining. When finished, will distribute to campus offices for comment and approval. Data Stewards Committee and Committee on Institutional Data will review, analyze, and fund approval. When approved, will be mandated as university practice.

Archives has examined, compared and mapped numerous metadata requirements/projects E.g., Australia. Formatted as a chart, with requirements mapped one-to-another, with similarities and differences noted.

Strength =record keeping, not as strong with other metadata requirements

3. Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in anyway to facilitate education? What's your preferred way to learn?

Educational challenge – high turnover among personnel in offices, as staff, analysts and designers come and go. Education has to be continuous, with monitoring of comings and goings of new people, messages continually relayed and renewed.

4. Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

IU has goal of managing digital resources across the campus. Archives involved, especially since 1995: NHPRC grant for developing a strategy for ER management. Phase II began with another NHPRC grant in 2000.

Developing partners, identifying constituencies that have something in common. Long process of negotiation. Strongest partners: internal auditors, common intersections of objectives, strategy to work together. Archives and auditors do audits as a team, effort is past experimental stage. Also working with data people. IU has data stewards, responsible for data in large area, e.g., student record keeping or HR. Stewards meet regularly, Phil is a member. Above this level is the Committee on Institutional Data. Made up of policy makers, Deans, Assoc VPs etc. They move policies/procedures through system to be approved.

Enterprise-wide initiatives include document management, records management and keeping in systems, high level planning in creating systems; preservation of digital objects; library community-metadata and digital library program and preservation.

Specific projects - Records Management workshops, one-on-one counseling, Archives' web page, development policies and procedures on regular basis. In the draft stage are policies on e-mail, scanning/imaging

What do the constituents know?

Understanding of metadata requirements for their area, including for metadata as relates to record keeping needs. May not understand what kind of applications for record keeping metadata.

XML

Archives knows little about XML, so not educating constituents about it. Constituents have some recognition that XML has potential for retrieval and access-digital lib community is an example. Among the Data Mgmt group, info mgmt, IT, XML not so well understood.

General need to make the jobs of records creators easier, not add on to current duties. Should choose wise examples for them to see the need for change.

Project

(Rosemary Pleva Flynn)

Portal with infrastructure-record keeping functionality: My IU is now called OneStart. Linked with Enterprise Development Environment (EDEN) as infrastructure. Need to replace enterprise-wide systems,. Question: how to provide access to systems from one central location. Answer led to portal concept. Meant to be environment for all HR and student information needs, employees, professors, students, one-stop-shopping.

Focus on work flow engine, routing of documents to follow formal review and approval process. Need to make sure requirements for record keeping and metadata built into work flow so tracking, and modifications would be noted in audit trail. Archives worked with developers of EDEN and portal, to learn what metadata is to them, what metadata is needed for the process. Now developing workflow engine, which won't automatically collect all record keeping information the system needs. Deal with record keeping at enterprise level. Product of seven

years of work. RFP notes that record keeping is a priority. Portal plans to use XML in data-exchange.

Conclusion

Archives has established partners and identified others to work with. Strong expertise in metadata. Knowledge about XML is biggest deficiency.

Strong interest in business process modeling for design of new systems. It's the roadmap to records - identifies where records occur, then assigns value, need to capture, what kind of metadata is appropriate. Springboard to all other discussions on record keeping. Would like to role business process modeling into whatever Archives does.

Delaware Public Archives

1. XML

How have you learned about XML?

Learning about XML in the Delaware Public Archives has been primarily limited to an ad-hoc learning style based primarily on on-line and vendor-supplied resources. Xml.com has been a primary source for educational information on XML. Several years ago senior DPA staff did attend briefings on the EAD.

Staff of DPA's sister agency, the Division of Libraries, have attended formal SGML, XML, and metadata training at GCA, OCLC, and DC6. Also, staff attended the first state GILS meeting in Olympia, Washington in 1999.

What XML-related projects are you working on?

The Delaware Government Information Center (and, by extension, the Delaware Public Archives) is currently involved in a number of projects that are XML-related. These projects include:

<p>Guide to the Holdings of DPA</p> <p>This collection of approximately 3500, structured WordPerfect files have been successfully converted to HTML. This collection will be available when a new DPA web site is released for public use at the end of March, 2000.</p> <p>Work is also underway to refine the templates used to produce the Guide to create well-formed XML utilizing the EAD tagset. A third phase of this effort will utilize the transformation capability of the Bluestone XML server to produce a "valid" EAD finding aid.</p>	<p>HTML Transit</p> <p>Transit is a powerful, template driven conversion tool that can generate both HTML and XML from many word processing and spreadsheet inputs.</p> <p>Transit supports custom tagging and the capability to toggle inclusion of any defined element in an output product. However, Transit will not restructure or transform documents.</p> <p>This product is also being considered for use as the initial formatting tool for an XML version of the Delaware Register. Staff of the GIC and the Archives are providing the legislative office compiling the Register with limited design and technical support.</p>
<p>Web-enabled GIS mapping</p> <p>The GIC has recently licensed ESRI's</p>	<p>ARC IMS</p> <p>This product utilizes XML for data</p>

<p>ArcIMS software to publish maps and geospatial data to the Web. The first datasets selected for publishing are a comprehensive set of state land use and transportation layers. Other data being considered include the point data for DPA's Historic Markers program, site profiles of the state's historically and archaeologically significant sites, and the boundary data for the state's new legislative districts.</p>	<p>exchange between its various query, processing and rendering services.</p> <p>The GIC is investigating a process to capture or extract these internal data streams for archival purposes.</p>
<p>Portal Development</p> <p>Delaware's new portal is scheduled for release the last week of March. It will utilize the Inktomi spider and search engine. The Government Information Center will coordinate the updates to the portal, and manage the system when it moves to an 'in house' service at the end of the present contract for remote hosting.</p>	<p>Ultraseek</p> <p>This Inktomi search engine provides the capability to index and search well-formed XML 'out of the box' and valid XML if a DTD is identified to the application.</p>
<p>On-line franchise tax</p> <p>The Division of Corporations has just released a web-based application enabling the online payment of corporate franchise taxes. DPA may have an opportunity to test the capture of the transaction record in an XML container.</p>	<p>Bluestone</p> <p>Licensed by the Department of State for its application in the e-commerce environment, Bluestone provides the Archives with an XML document transformation server and a web-database query and presentation engine.</p>

2. Metadata

How would you define metadata?

In 1998, DPA issued the Model Guidelines for Electronic Records, which provided a rudimentary definition of metadata. The guidelines also made mention of a metadata model to be adopted by DPA in conjunction with further work on electronic records. To date, that work has not been accomplished. DPA did not accept the University of Pittsburgh metadata model (for business acceptable communications), but only because the scope of the initial project did not call for a validation/adoption process.

What standards are you using or learning about? What project are you applying them to?

DPA had initial discussion during FY 00 with the Office of Information Services about the adoption of a metadata model for record keeping, in conjunction with a data warehousing initiative. At the time, the Australian Record keeping Metadata model was used as the starting point of discussions. However, no progress has been made on this in the past 10 months.

3. Education and communication

Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)?

DPA offers a series of workshops and classes on a regular basis to records creators. These workshops are offered through our State Personnel Office and are taught at sites across the state, and are available to all state employees. DPA recently initiated a new course—“Managing Records on Your Personal Computer”—as part of its offerings. In addition, DPA co-sponsored with the Office of Information Services a one-day seminar on electronic record keeping issues in April 2000.

DPA has used its web site for dissemination of records management products with mixed results. All policy and procedural manuals in print are available on-line; however, initial findings suggest that printed (hard) copies remain the medium of choice for the majority of our constituents.

DPA does have access to the use of a state-of-the-art computer-based learning lab at our Division of Libraries.

4. Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

DPA's primary constituency for record keeping issues remain the agency records officers who have jurisdiction over the management of paper records in agencies. As DPA ventured into electronic records management issues, we quickly learned that there was a disconnect apparent between our traditional agency records officers and the information technology-side managers. In many events, traditional AROs had little, if any, contact with IT-side resources. As part of the April 2000, OIS and DPA sought to bring together these disparate resources. The one-day seminar resulted in 110 participants from more than 30 agencies.

The critical need, however, is to link DPA's interest in electronic record keeping issues with other resources within an agency. Specifically, the issue of electronic record keeping needs to be addressed by resources responsible for legal issues, and content issues in the agency. There is an emerging constituency in agencies who are in need of electronic record keeping solutions and view the solution as a policy issue, and not a strictly information technology issue.

Ohio Historical Society

1. XML

How have you learned about XML? What XML-related projects are you working on?

- *Introduction to XML*, Microcenter Class
- State Technology Briefing -- Gartner Group
- Various publications

The Joint Electronic Records Repository Initiative and the Electronic Publications Pilot Project will utilize XML. As these projects are really in the planning stages it is hard to say anything specific about how XML will be used.

2. Metadata

How would you define metadata? What standards are you using or learning about? What projects are you applying them to?

Data about data

EAD finding aids

We are working toward creating our finding aids in EAD and linking to them directly from the MARC record in the OPAC.

Dublin Core/GILS

The Electronic Publications Pilot Project is an ongoing project to identify software that will capture state publications on the web for permanent storage. In conjunction with this the project team is investigating using Dublin Core metadata or possibly using a adapted GILS metadata schema.

Australian Recordkeeping Metadata

As part of the Joint Electronic Records Repository Initiative we will be studying this metadata information for possibly use in the long term preservation of electronic records.

GILS

The Ohio GILS project makes state agencies retention schedules available online with descriptive information about the records described in those schedules. The metadata set used is based on the GILS schema.

3. Education and communication

Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in any way to facilitate education? What's your preferred way to learn?

Local Government Records Program (LGRP)

- LGR: Just the Basics -- an introductory workshop on records management presented in traditional workshop format and on the web as a Powerpoint with voice over. Approximately ten workshops are held throughout the state annually.
- Managing Electronic Records -- an introductory workshop on electronic records issues and scheduling held throughout the state
- Powerpoints from various presentations made available on the Web for download
- Handbook -- published online and hardcopy. Introductory records management information
- Manuals -- published online and hardcopy. Suggested retention periods for local government entities
- Electronic records guidelines -- email, imaging, records management guidelines available online and in hardcopy

State Government

- Electronic Records Committee policy and guidelines -- email, imaging, records management, TIS (in progress), file management (in progress). The ERC also uses a listserv to facilitate communication and distribute information.
<http://www.ohiojunction.net/erc>
- At the state level we do a substantial amount of work educating agency personnel on a case by case basis as they approach us with questions and problems.

Generally we rely heavily on the web to disseminate resources and to provide training. All resources are provided online and our hardcopy distribution is consistently declining. We are looking to providing more interactive and voice over training resources via the web.

Learning styles vary. We have found that at the state government level users want everything on-line and straightforward -- lots of bullets and easy to read text and we are trying to create our new resources in this style. Ideally it seems like you have to provide educational materials that suit all the different learning styles -- text, outlines, graphics, hands-on, online, individualized meetings and larger meetings.

4. Constituencies

Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

We work closely with several members of the Ohio Electronic Records Committee which include records managers, attorneys, information technology specialists, and archivists. We also work with the State Library of Ohio, the Department of Administrative Services, and the Ohio Supercomputer Center on the JERRI project. The JERRI partners combine the expertise of archivists, librarians, records managers, legal/policy experts, and technology specialists in an effort to bring all the appropriate specializations together to work on the issues surrounding electronic records.

At the state level there is a monthly IT Briefing usually given by Gartner Group or other IT consultants. We have given one of these briefings in the past and view them as a possible delivery system for XML education.

Kentucky Department for Libraries and Archives

1. XML

How have you learned about XML? What XML-related projects are you working on?

Articles, books, online resources:

- ElementK instructor led online course "Introduction to XML" - currently half way through it.
- MN State Archives XML web site <http://www.mnhs.org/preserve/records/xml.html>

Classes, workshops, presentations:

- Richard Marciano presentation on SDSC project at MER in Chicago Oct. 2000

Projects:

- Document Management Digitization System proposed for Kentucky Department for Libraries and Archives
- Potential XML Project with Secretary of State Office, supported by R&D grant from the Governor's Office for Technology

2. Metadata

How would you define metadata? What standards are you using or learning about? What projects are you applying them to?

Definition: Data about data; information used to facilitate the location of, access to, management of, or evaluation of other information (see web site - <http://www.mnhs.org/preserve/records/metadata.html>)

Standards and projects:

- EAD finding aids in connection with Kentuckiana Virtual Library Project (Multi-Institution Project. (see <http://www.kyvl.org/kentuckiana/digilibcoll/digilibcoll.shtml>)
- Dublin Core (using Tag Gen for web page metadata) as part of GIS metadata at the Kentucky Department for Libraries and Archives. (see the Department's GILS project, Findit! Kentucky at <http://www.finditkentucky.net>)
- Ongoing Web Standards Committee work with an opportunity to revise statewide architecture standards for metadata.

3. Education and communication

Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in anyway to facilitate education? What's your preferred way to learn?

KDLA: We regularly conduct workshops connected with our records management program. We also have conducted meta tagging training workshops connected with our Findit-Kentucky project. The Department has also held focus groups meetings soliciting opinion on electronic and non-electronic recordkeeping practices. State and local government officials from a variety of institutions and people at various levels of the governmental hierarchy participated in these meetings

We rely heavily on the Web to present materials to the public .

GOT: The Governor's Office for Technology is in charge of IT training for state agencies. The division that conducts that training is in the Office of Policy and Customer Relations, the same unit that develops IT standards and policies for the executive branch. We believe that where XML is relevant to agency needs, there will be a necessity to develop supportive training curricula.

4. Constituencies

Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

KDLA: We try to create tools that anyone working with electronic records can use (see http://www.kdla.net/pubrec/Recordkeeping_and_new_technologi.htm), but we work closely with a smaller group of CIOs, records managers, and data administrators on a continuing basis. We are currently undergoing a review of our electronic records program and evaluation of our capacity to promote archival services, in the context of a state wide emphasis on doing more state government business electronically. (see Enterprise Electronic Records Program: Assessment and Report - <http://www.kdla.net/pubrec/Electronic%20Records%20Program%20Assessment%20Update.htm>)

GOT: As the foregoing suggests, a primary focus of the Governor's Office for Technology is to enhance service provision through increased reliance on egovernment transactions, and hence, electronic records creation and management become more and more critical elements to the success of that strategy. We believe XML will increasingly be used by a variety of the software components we and agencies will purchase or develop to facilitate ebusiness, and we want to be developing any appropriate standards that enhance our capacity to conduct our business. It would be accurate to say we are much more front-end focused than KDLA, but we hypothesize that getting it right at that point will create significant opportunities to work out the end-of-life-cycle management issues.

Smithsonian Institution Archives

Smithsonian program is different – younger, less focused, has attempted to do it all. XML and metadata are opportunities to focus and move forward.

E-mail decentralized, as well as collection management systems
10 archives

SIA archives and records management programs, Collections management policy (all at institution)-stronger interest in collections-related info

Digitization mantra at Smithsonian, 1998-2000

1. How have you learned about XML? What XML-related projects are you working on?

Educate themselves (Electronic records issues in general: Slavin and Horton presentations)

No implementation

2. How would you define metadata? What standards are you using or learning about? What projects are you applying them to?

Dublin Core testing (CIMI-Consortium for the Interchange of Museum Information)-searching collections at high-level, objects, books, and archives over the web

Lead to Collections Information Public Access System (CIPAS)

Discussed in record keeping advice
Incorporated into all accessioning planning

3. Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in anyway to facilitate education? What's your preferred way to learn?

Heavy web orientation-decentralized by museum, provide content
<http://www.si.edu>

Developing written guidelines for broad use by Smithsonian staff
Draft brochure on rm and edm systems
Draft brochure on e-mail
Websites-record and brochure
Financial systems-report on record keeping capabilities of PeopleSoft
Smaller more general brochures-hope will be more widely read

Other guidance

Doculabs-records management concerns, recommendation is not to buy rm/edms software, but SIA has brochure ready to go, asking Gartner for a second opinion.

SIA website is being expanded. Participating in new employee orientation to get ER message out. Publishing in HR magazine (e-mail netiquette), e.g., proper and improper times to use e-mail to

Beginning use of websites, handouts, workshops

4. Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

Office of Information Technology (OIT)

SIA controls ER but OIT has responsibility of preservation. Example of system: correspondence tracking system in Office of the Secretary

Smithsonian Institution offices

OIT, general counsel, Webmasters, data managers, collection managers

Constituents systems: e-mail, web systems, document management systems (Secretary's Office, Office of General Counsel, Comptroller and other financial and contracting offices), collections systems, legacy systems

XML-capture legacy data/systems in XML format so available to everyone in common format??
Deal with so many organization entities and types of data

Technological questions and policy questions

Questions

PB: Is SIA currently using TRIM?

EH: No, technical and intellectual problems with implementation

Minnesota Historical Society

1. XML

How have you learned about XML? What XML-related projects are you working on?

Articles, books, online resources:

- . Simon St. Laurent, *XML: A Primer*
- . MN State Archives XML web site <http://www.mnhs.org/preserve/records/xml.html>
- . Periodicals (DM Magazine, Infoworld, KM Week, Civic.com, Government Technology etc.)

Classes, workshops, presentations:

- . March 2000 workshop in St. Paul – Richard Marciano and Chaitan Baru
- . Alto Consulting and Training course (three days)
- . MN Science Museum course (two days)
- . Vendor presentations (Datachannel)
- . User Groups (MN SGML-XML Forum)

Projects:

- . Educating Archivists and their Constituencies

2. Metadata

How would you define metadata? What standards are you using or learning about? What projects are you applying them to?

Definition:

Data about data; information used to facilitate the location of, access to, management of, or evaluation of other information (see web site - <http://www.mnhs.org/preserve/records/metadata.html>)

Standards and projects:

- . EAD finding aids
- . Dublin Core (using Tag Gen for web page metadata. See Bridges web site - <http://bridges.state.mn.us/>) Also connected to GILS
- . GIS metadata (FGDC and MN flavors – see <http://www.lmic.state.mn.us>)
- . Australian Recordkeeping Metadata (Metadata Development Comm – see <http://www.mnhs.org/preserve/records/metadev.html>)
- . ISO 11179 data elements
- . Various studies (<http://www.mnhs.org/preserve/records/metadata.html>)

3. Education and communication

Do you offer regularly any educational opportunities to records creators (e.g., workshops, classes, manuals)? Do you use the Web in anyway to facilitate education? What's your preferred way to learn?

We sponsor and provide presentations on our own and in collaboration with a wide variety of groups (records management, information technology, government, vendors, universities). Most of these go online in some form or another (e.g., as lists of resources, as grist for manuals and guidelines, or just PowerPoint files). We have also hosted focus groups on many issues (metadata, data warehouses, electronic records).

We rely heavily on the Web. We establish pages for workgroups in order to facilitate information sharing and coordinate efforts. We provide all resources online, usually in both HTML and PDF formats. The Web allows for quick and easy updates, at a very low cost, with links to our partners and other resources. All indications are that about 100% of our primary audiences have access to the Web and that they like using this as the medium of communication.

Everyone's learning style is different, but there's a general consensus that hands-on learning is best. Having an instructor or some other means to get answers to questions quickly and easily is necessary too. Additional resources for further learning after a class is good. In all, a variety of complementary methods are ideal. We take advantage of university, vendor, and government class offerings as necessary.

4. Constituencies

Could you describe your primary electronic records and information technology constituencies? How do their interests and needs correspond or connect to yours?

We try to create tools that anyone working with electronic records can use (e.g., trustworthy information systems handbook), but we work closely with a smaller group of CIOs, records managers, and data administrators on a continuing basis. These people for the most part form part of formal and informal organizations: the Information Policy Council, the Data Issues Group, the Government Records Information Network, etc. We also work with these people and others on projects, such as the electronic real estate recording task force, record keeping metadata development committee, a number of GIS efforts, and so on. The smaller groups are largely selected or self-selected, based on their interest in information technology and record keeping. They form a kind of readily available focus group and proving ground that keeps us practically oriented and productive. It's a mutual educational process: they learn about archival and record keeping issues and we learn about information technology. We look quite actively for areas of overlapping concerns and try to explore them. Lately, we are concentrating on metadata and XML, as tools that will be of mutual benefit to everyone.