Hello, I'm Melanie Feltner-Reichert, director of Digital Library Initiatives at the University of Tennessee. My colleague, Linda Phillips, is going to set the context for Metadata Plus, and I'll pick up in the middle when we get to the demonstration. Good morning, I'm Linda Phillips, head of Scholarly Communications at UT; our colleague, Marie Garrett sends you greetings from Knoxville.
This presentation is about discovering locally created digital content. Our premise is: Most librarians understand that metadata facilitates discovery of digital content, but most of us cannot envision how it happens. One benefit to understanding metadata’s inner workings is to be convincing when we encourage authors to entrust digital content to library collections.
To show how metadata works, we're going to define it, introduce you to the de facto standard for metadata creation—the Open Archives Initiative—and demonstrate the steps to create metadata that's harvestable. Because metadata isn't the only way to discover locally created content, we'll also mention some of the traditional bibliographic access methods that expose digital content.
Academic library digital programs are flourishing. Digitization of unique content brings special collections to the desktops of scholars and other information seekers from kindergarten to life-long learners.
Standing on the shoulders of library digitization experiences, new forms of scholarly publishing are emerging—for example, digital imprints such as the University of Tennessee Newfound Press,
Institutional repositories…
Subject guides created by librarians,
And open access journals.
Although several options are available to discover online resources, faculty and students are likely to use web search engines first.
For more in-depth research, scholars rely on a variety of tools, many online, that provide access via bibliographic conventions developed for access to print. For example, someone seeking information about government policy toward African American civil rights might well look in the library catalog and in appropriate databases. You probably recognize the box on the right as an excerpt from a MARC record. The MARC cataloging record makes our Newfound Press online book discoverable through the local catalog and WorldCat; a metadata record enables its discovery via internet harvesters.
The OAIster database is a union catalog of digital resources whose metadata conforms to the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH for short). Although metadata as a concept has become increasingly familiar to scholars, many are unaware that OAIster can be used much like a commercial database to search and retrieve digital resources. OAIster is just one of many metadata harvesters.
We’re now going to enable you to visualize metadata in action. Melanie: I will now address the role of metadata in making locally created content discoverable. To do that, we’ll look at what metadata is, the features of OAI-compliant metadata records, and show you how to participate in the Open Archives Initiative. Metadata: let’s watch it work.
Melanie’s first slide: But - a short aside. Before we launch into the definition of metadata, I'd like to mention an exciting development in the world of digital resources. Here we see an example metadata at work. Linda just mentioned OAISTER, a union catalog of digital resources maintained by the University of Michigan. You could say OAISTER is the WorldCat of the digital world. Interestingly, just in January, OCLC and UMich developed a partnership with an end goal of transferring responsibility for maintaining this digital union catalog to OCLC. .... With this development, we see OAlster becoming a more significant as a database. This represents a turning point in providing access to digital content -- offering a promise of sustainability in providing access to digital objects. And I point this out because our entire talk is about making digital content discoverable. So now let's look at how that works, beginning with the role of metadata.
So, first, what is metadata. Here’s a simple, straightforward definition created by Priscilla Caplan. Metadata is …… Sound familiar? we’re all familiar with the role of MARC records in capturing structured information about resources, specifically descriptive access points, helping users find content geared to their information needs. So, first, what is metadata. Here's a simple, straightforward definition created by Priscilla Caplan. Metadata is structured information about an information resource of any media type or format”. This definition is broad enough to encompass all forms of metadata, including descriptive, technical, administrative and structural. For the purposes of end-user discovery of information resources, descriptive metadata is the key. MARC is perhaps the descriptive metadata format with which librarians are most familiar. We’re all acquainted with the role of MARC records in capturing structured information about resources, specifically descriptive access points, helping users find content geared to their information needs.
Here we see two descriptive records for a single resource, one in MARC, and another in Dublin Core. Each record describes a digital resource published by our library’s electronic press. Notice that each record shares similar descriptive access points, such as title, author, subject. However, the Dublin Core record has a few fields that are organized and presented a bit differently: These fields are highlighted in red. Unique identifiers are captured in repeated "identifier" fields. The first field is the unique file name assigned to the object. The second is the persistent URL where the resource resides. Format and type are also structured differently to conform to best practices for Open Archives Initiatives records. The format field, for instance, is expressed in a way using the file type controlled vocabulary required for OAI-compliant Dublin Core. So, as you can see, it is not such a big leap to move from the familiar MARC format to a Dublin Core format, but certain tweaks must be made to the structure of the information to conform to OAI best practices.
Now let's look a little deeper into OAI compliance. This slide takes us "behind the scenes" of the same Dublin Core record just shown. This is the same descriptive content -- however, it is now expressed in valid XML (extensible markup language). This standardized encoding is a critical part of what makes the record harvestable for OAI service providers such as OAIster.
So how does OAI-complaint encoding of your records make your locally created content easier to find. This diagram illustrates how the harvesting process works - bringing information to the user! OAI complaint Dublin Core records for digital collections are placed in an online repository by the holding institution. The data is then picked up by OAI harvesters maintained by service providers, then delivered to the end user in a new database of aggregated records searchable through an interface like OAIster. Databases like OAIster provide "one stop shopping for users interested in useful, academically-oriented digital resources. [they] gather all potential digital resources out there in an effort to build a comprehensive digital union catalog" The service also allows users to quickly access the information object itself, not just its record. Each record links directly to the resource online as illustrated in the Dublin Core record we viewed a moment ago. This eliminates frustration and dead ends. The user finds a record describing an object of use and is then immediately able to retrieve the object and put it to use in their research.
We've mentioned OAISter a lot; however NDLTD is another example of a service built on OAI records.
Here’s an example of one of our university’s dissertations indexed in NDLTD. Because we share our records via the OAI protocol, the NDLTD service harvests our records from the repository we set up. They then index the records in their database, making it easier for researchers to find dissertations and theses created at our institution.
Here we see the digital object listed in the NDLTD search results on the previous slide, and its OAI-complaint record. We created an OAI-compliant record for the object, placed it in our repository and NDLTD harvested it! This illustrates the benefits of participating in the OAI initiative: Library records available to all OAI harvesters and service providers, the chances of content discovery increase, and the visibility of the author's work is thus enhanced.

>
Participate in OAI-PMH

Step 1
*Decide on static vs. dynamic*

**Static**
- Best for collections that are:
  - Small (1-5,000 items)
  - Established
  - Have little technical support

**Dynamic**
- Best for collections that are:
  - Large (More than 5,000 items)
  - Growing
  - Have significant technical support

Read slide. Step 1: Decide whether to set up a static or dynamic repository.
Step 2. The institution must set up a repository. A repository is simply a server that hosts OAI-compliant records for harvesting.
Step 3

Create records

- Combine simple Dublin Core records for one collection into a single XML file.

- Validate the file according to instructions on the OAI site.

- Save the file according to OAI naming conventions.

Step 3. To create records to place in your repository, combine all simple Dublin Core records for one collection into a single XML file. Validate the file according to instructions on the OAI site and save the file according to OAI naming conventions.
The final step is to Register as data provider with the initiative. Submit and validate your URL; enter your institution's name, contact information, and the version of OAI-PMH used in the repository. It's as simple as that!
The reward . . .

. . . for this investment is global visibility of locally created content via numerous gateways and search engines!

The reward for your OAI participation is both global visibility of the digital content generated at your institution, and worldwide access to the information. Although the guts of our presentation has focused on metadata, many traditional methods for making content visible in the print environment work quite well for locally created digital content. Linda will wrap up with a summary of strategies we're using at the University of Tennessee.
This is the Plus section. At UT we’re adding local content to our collections in the form of peer reviewed publications, digital repository contributions, and open access scholarship that’s freely available on the web. In the context of our digital imprint, Newfound Press, we've given considerable thought recently to the ways prospective users will discover our publications. These strategies are among those we’re pursuing—we’re cataloging the new content, purchasing ISBN and securing ISSN numbers, soliciting book reviews, and trying to get into Amazon.com.
Local cataloging ensures inclusion of a title in WorldCat. And, our partnership with University of Tennessee Press for print-on-demand provides more visibility, as well as an Amazon.com listing. A current challenge is getting the book reviewed, because traditional journal editors decide among the plethora of print books they receive which few will be reviewed. Newfound Press simply doesn’t have the funding to give away more than a copy to the author, and besides, we’re a DIGITAL press. I asked our approval vendor about potential benefits to them to distribute our books which are free online and come with stellar cataloging, but haven’t yet received a reply—I hope this idea is only slightly ahead of its time.
To make locally published journals discoverable Newfound Press applies for ISSNs. MARC records and OAI-PMH make the title discoverable through library catalogs and internet harvesters. We have not yet attempted, but intend to approach appropriate indexing and abstracting services, as well as Ulrichsweb to include our journals in their directory. Further, in the same spirit as approaching an approval vendor, we could offer free content to journal aggregators. Another goal is to explore the process for securing Digital Object Identifiers. We are a few years away from approaching Journal Citation Reports to seek inclusion, but this would be a logical next step.
In this digital age, librarians have increasingly frequent conversations with faculty and others about placing local content online in the library. We advertise libraries as the location of choice for assuring discovery, access, and preservation. Metadata is our most current and, possibly, most powerful discovery tool. Some of our traditional strategies for promoting discovery also scale well in the digital environment. We hope this morning’s demonstration will help you assure the discovery of your locally created content. We welcome your questions.