no new money coming in

amateur error

largest distributed academic digital library in the world

RePEc

does not represent an official statement from anybody

please interpret me at any time

assumes some level of familiarity with RePEc

about this talk

2001-09-17

http://opend.a.org/home/Knichel

Long Island University

Patrick School of Library and Information Sciences

Thomas Knichel

RePEc and OA and AMF
stimulate discussions
• call for a meeting in January for “proposals” prepared to

• emphasis on a pragmatic level of interoperability

archived literature
- to work towards achieving a universal service for authors.
“the purpose of this call is the mobilisation of a core group
Ginsburg, Luce and van de Sompele (1999)

Open Archive Initiative History

Open Archive Initiative (OAI) and AMF "Academic Metadata"

- worked system that is not taken up in other disciplines
- no long-running for development work
- loose coupling of RIF
- template syntax of RIF
- no relational checking
- no central registration of service

Weaknesses of RIF

starts to make inroads in the peer-review process

• separate log treatment

– separate institutional registration
– separate personal registration

References

high degree of labour division

pioneered distinction between provision and usage of digital

RePEc Statistics
In brief:

- "Record" is metadata like in full-text archives.
- Open-archival archives are like a record archive.
- The "managed" or formal-print archives: not papers on the web.
- Basic concepts

Sponsored by CILR, LANL, and SPARC

Chair: Luft and Waters

NDT, REPE, SLAC/SPIRES, and others

Representatives of axiety, coffee, HEP, and others

Santa Fe Meeting 1999–20

This work was done to prepare the Santa Fe meeting, and pub-

- Lack of incremental harvesting
- Under legal terms and conditions
- No uniform identifier structure
- Poor metadata

ability between report initiative and report service. They find that the main problems of Internet-
argval is the value of the attribute
argname is the name of the attribute
vername is the name of the verb

at location of the OAI-PMH protocol as registered

baseurl=\&argname\&argname\&argname\&argname

where Compliant archive responds to the request "verb" encoded

OAI Technical model: OAI protocol for Metadata Harvesting

Each record has a datstamp:

Records may be grouped into sets:

Not all records need to exist in all formats

Unqualified DC required metadata for each record

"Native" metadata format

of any number of metadata formats.
The main extension that OAI offers over RePec is the presence

Open Archives metadata model

much younger:

OAI about 30 data providers and 3 service providers, but it is

Many services (service providers)

— One logical database (in the RePec case)

— Many archives (data providers)

Separation between data providers and service providers

Inspired by RePec initiative

Open Archives business model
OAI PMH: verbs

- Identify
- ListSets
- ListMetadataFormats
- GetRecord
- ListRecords

OAI PMH: Sets

- ONLY for a local community to implement selective harvesting
- there can be zero or more sets in an archive
- records can exist at interior nodes in the set hierarchy
- asking for records in a set, returns records in the set and in all its subsets.
- currently not used in RePEc OAI archive, but could be applied to archives and series

OAI PMH: XML Schema usage

All responses to OAI PMH requests are controlled by XML Schema instances.
This includes the header of the response as well as the metadata instances that the header may include.
Note that version 1 of the OA PMH protocol had to be revised to version 1.1 in July 2001 following a change in the XML Schema specification.
RePEc OAI PMD implementation

- spare-time job by Thomas Kirchel
- runs on arcana, the first machine that WoPMD bought with JISC money
- move to carnation.iert-hlt-u.ac.jp delayed by hardware problems

Morton protocol

A protocol to store files on a disk that come from RePEc but are stored for delivery of OAI PMD data provider. Idea: make everything available via ftp, too, for us oldies.

etc/ files for delivery of general metadata information
lib/ files of metadata records
var/ files that have the identifier per timestamp
bin/ implementation software
directory etc/

There is a written-down version of this, but not up-to-date.

- Identify
- ListMetadataFormats
- ListSets

the latter one will need moving once we represent RePEc series and archives as sets.
implement the OAI PMH responses.

Very I/O intensive, very slow.

delivers from the installed file, it installs the temporary file.

in 'lib', it creates a temporary file, and it the temporary file.

the contents of REG file, for files in 'lib', anew, for files

OLDC, the unhandled DC is returned by the OAI if build

redacted a perl script that translates REDIF to AMP and

software to run all this

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl

directory perl
More formal semantics at http://www.openlib.org/doc/epubxml
Design technical at http://www.openlib.org/home/technical/technical.html
Krikel and Simmone M. Warren for all.

Creation follows decision at workshop of the OA1 Technical Com-

use standard XML validation and API tools

more general

simpler to create

Am: to create a successful to RDF in XML.

AMF, the Academic Metadata Format

self-organizatoin to the format.

Thus we need to use the OA1 to pull our agendas of academic

conversion of the OA1 activities program.

standing communication. Continuing support of this work remains a

access to open archives as a means of increasing the availability of

The OA1 archive initiative has its roots in an effort to maintain

movements. On the http://www.openarchives.org

to the origin of the electronic theses and dissertation (ETD)

This culture overlooks that the main players are still nested

format transported. This is of little use for us.

Now all it has produced is a DC and maybe other metadata

entity archives as a means to promote their universal use.

OAI set out at the beginning to work on the interoperability of

OAI culture

May also be useful to feed other AMF-based data into RDF.

Gateway will be maintained centrally. That gateway, however,

started.

protocol with OAI PMH so all RDF archives. That is a non-

near future, it will be prohibitively costly to replace the Guardian

OAI PMH much more involved then Guardian protocol. In the
still in need of a bit of TLC

XML Schema based

AMF use standard XML validation

could be added.
Currently only an resource type, "text", but software or others

collection noun
It allows for arbitrary grouping of resource records in the

AMF is more general than RDF

The documentation is only 7 pages long.

There are "verbs" that allow to relate nouns.

Each noun admits XML child elements called "attribute", that
called "nouns", Each noun is represented by an XML element.

text - collection - person - organization.
There are only four record types in AMF

AMF is simpler to create than RDF