Open Archiving
@
Virginia Tech
DLRL

Hussein Suleman
http://www.dlib.vt.edu
Why are we involved?

- Networked Digital Library of Theses and Dissertations (NDLTD)
- MARIAN Information Retrieval System
- Computer Science Teaching Center (CSTC)
- PetaPlex
- Web Characterization Repository
- VT-Imagebase, CDDC, etc.
Alpha Test Collection

◆ VT ETD Collection
  ■ Based at university library
  ■ Accessible through remote OAI interface
  ■ Approximately 3000 records
  ■ Dynamic

◆ NDLTD
  ■ Distributed digital library
  ■ Willingness of participants?

http://smp2.cc.vt.edu/cgi-bin/hussein/vtetd/OAI1.0
Issues in OAI-Compliancy

- Cleaning of data
- Rights management and ownership
- USMARC, rfc1807 and DC
- Resumption tokens
- Existing protocols and layering

http://smp2.cc.vt.edu/cgi-bin/hussein/vtetd/OAI1.0
Technical Requirements

- Implemented in Perl for flexibility and platform-independence
- Original implementation took 3 hours, after a first reading of protocol!

http://smp2.cc.vt.edu/cgi-bin/hussein/vtetd/OAI1.0
Repository Explorer

- Need consistency for NDLTD
- Online vs Offline
- Protocol and XSD testing
- Interactive browsing
- Automatic test suite
- List of demonstration archives

http://purl.org/net/oai_explorer
Open Archives Initiative Repository Explorer v1.0

Please enter the URL to the OAI interface (everything before the ?)

Start Browsing

Some predefined archives:

American Memory [LoC] (archive details)
Virginia Tech Thesis/Dissertation Collection [VT] (archive details)
Digital Cultural Heritage Community [UIUC] (archive details)
LDC [Open Language Archives] (archive details)
Perseus [Open Language Archives] (archive details)
Lacite [Open Language Archives] (archive details)
RDW [Resource Discovery Network] (archive details)
LTRS [NASA] (archive details)
NACA [NASA] (archive details)
arXiv [LANL] (archive details)

Test and Add an archive to this list

http://purl.org/net/oai_explorer
Open Archives Initiative Repository Explorer

http://purl.org/net/oai_explorer
Open Archives Initiative Repository Explorer v1.0

http://smp2.cc.vt.edu/cgi-bin/hussein/vtetd/OAI1.0?verb=Identify

Archive details: http://etd.vt.edu

Archive Self-Description

<table>
<thead>
<tr>
<th>Repository Name</th>
<th>Virginia Tech Electronic Thesis and Dissertation Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base URL</td>
<td><a href="http://researchsmp2.cc.vt.edu:80/cgi-bin/hussein/vtetd/OAI1.0">http://researchsmp2.cc.vt.edu:80/cgi-bin/hussein/vtetd/OAI1.0</a></td>
</tr>
<tr>
<td>Protocol Version</td>
<td>1.0</td>
</tr>
<tr>
<td>Admin Email</td>
<td><a href="mailto:hussein@vt.edu">mailto:hussein@vt.edu</a></td>
</tr>
<tr>
<td>Description</td>
<td>oai-identifier:</td>
</tr>
</tbody>
</table>

http://purl.org/net/oai_explorer
List of Metadata Formats

Click on the link to view schema

Prefix=[oai_rfc1807]
Schema=[http://www.openarchives.org/OAI/rfc1807.xsd]

Prefix=[oai_marc]
NameSpace=[http://www.openarchives.org/OAI/oai_marc]
Schema=[http://www.openarchives.org/OAI/oai_marc.xsd]

Prefix=[oai_dc]
NameSpace=[http://purl.org/dc/elements/1.1/]
Schema=[http://www.openarchives.org/OAI/dc.xsd]
List of Fields

Header:
Identifier : oai:VTED:etd-3345131939761081
Datestamp : 1997-03-31

Metadata:
oai_dc:
  title: Conceptual Development and Empirical Testing of an Outdoor RecreationExperience
  creator: Walker, Gordon James
  subject: Forestry
  description: This dissertation examines four issues, including: (a) whether outdoor
  contributor: Daniel R. Williams
  contributor: K. Jill Kiecolt
  contributor: Bradley R. Hertel
  contributor: Joseph W. Roggenbuck
  contributor: R. Bruce Hull
  publisher: Virginia Polytechnic Institute and State University
  date: 1997-03-31
  type: text
  format: application/pdf
  identifier: http://scholar.lib.vt.edu/theses/available/etd-3345131939761081/
  language: en-us
  rights: unrestricted

http://purl.org/net/oai_explorer
<table>
<thead>
<tr>
<th>Verbs</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identify</td>
<td>from (YYYY-MM-DD):</td>
</tr>
<tr>
<td></td>
<td>until (YYYY-MM-DD):</td>
</tr>
<tr>
<td>List Metadata Formats</td>
<td>metadataPrefix:</td>
</tr>
<tr>
<td>List Sets</td>
<td></td>
</tr>
<tr>
<td>List Identifiers</td>
<td>id: [oai:VTED:etd-334513]</td>
</tr>
<tr>
<td>List Records</td>
<td></td>
</tr>
<tr>
<td>Get Record</td>
<td></td>
</tr>
<tr>
<td>Parsed</td>
<td></td>
</tr>
<tr>
<td>Raw XML</td>
<td></td>
</tr>
</tbody>
</table>

Request URL: http://researchsm2.cc.vt.edu:80/cgi-bin/hussein/VTED/OAI1.0?verb=ListIdentifiers
Response Date: 2001-01-22T20:07:31-05:00

http://purl.org/net/oai_explorer
<?xml version="1.0" encoding="UTF-8"?>
<GetRecord xmlns="http://www.openarchives.org/OAI/1.0/OAI_GetRecord" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/1.0/">
<header>
<identifier,oai:VTED:etd-3345131939761081</identifier>
<datestamp>1997-03-31</datestamp>
</header>
<metadata>
<title>Conceptual Development and Empirical Testing of an Outdoor Recreation Experience Model</title>
<creator>Walker, Gordon James</creator>
<subject>Forestry</subject>
<description>This dissertation examines four issues, including: (a) whether outdoor recreation is a viable alternative to traditional education for the physically handicapped, (b) conceptualization of an outdoor recreation experience model, (c) the development of an empirical test of the model, and (d) the evaluation of the model using existing facilities and materials.</description>
<format>application/pdf</format>
<identifier>http://scholar.lib.vt.edu/theses/available/etd-3345131939761081/</identifier>
</oai_dc>
</metadata>
</GetRecord>

http://purl.org/net/oai_explorer
Open Archives Initiative Repository Explorer v1.0

http://smp2.cc.vt.edu/cgi-bin/hussein/vted/OAI0.2?verb=ListIdentifiers

Archive details: http://vted.vt.edu

List of Records

Select a link to view more information

Error: Illegal tag found

Resume from [!!!100]

Request URL: http://researchsmp2.cc.vt.edu:80/cgi-bin/hussein/vted/OAI0.2?verb=ListIdentifiers
Response Date: 2001-01-22T20:15:40-05:00

http://purl.org/net/oai_explorer
Ongoing Projects

- MARIAN OAI input/output filters
- oai_marc definition, encoding and conversion
- Perl, JAVA codebase
- NDLTD Union Collection
  - VTLS
  - PhysDis
Ongoing Projects 2

- VT Digital Library
  - Central digital document archive
- CSTC and iLumina
  - Distributed DL for educational material
Vision

Lots and lots of little archives all over the world

OAI

International ETD Archive

International Teaching Resources Archive

Etc ...