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

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)
Lacito [Open Language Archives] (archive details)
RDWN [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
v1.0

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

Archive details: http://vtetd.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_rfc:1807]
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/OAI1/dc.xsd]
### Open Archives Initiative Repository Explorer 1.0 - Netscape

**Request URL**: http://research.mp2.cc.vt.edu:80/cgi-bin/hussein/vtetd/OAI1.0?verb=ListIdentifiers  
**Response Date**: 2001-01-22T20:07:31-05:00

#### Verbs
- Identify
- List Metadata Formats
- List Sets
- List Identifiers
- List Records
- Get Record

#### Options
- from (YYYY-MM-DD): [ ]
- until (YYYY-MM-DD): [ ]
- metadataPrefix: oai_dc
- identifier: oai:VTED:etd-334513
- set: [ ]
- resumptionToken: [ ]

---

Send all comments to hussein@vt.edu --- Digital Libraries Research Laboratory@Virginia Tech

---

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/OAI_Metadata">
<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 experience elements are consistently identified by participants; (b) the extent to which outdoor recreation experience elements are consistently identified by different consumer populations; (c) the extent to which outdoor recreation experience elements are consistently identified by the same consumer population over time; and (d) the extent to which outdoor recreation experience elements are consistently identified by the same consumer population under different conditions. The results of this investigation provide important insights into the nature of outdoor recreation experiences and the ways in which these experiences influence consumer behavior. The findings also have implications for the design and implementation of outdoor recreation programs and policies.</description>
<contributor>Daniel R. Williams</contributor>
<contributor>K. Jill Kleco</contributor>
<contributor>Bradley R. Hertel</contributor>
<contributor>Joseph W. Roggenbuck</contributor>
<contributor>R. Bruce Hull</contributor>
<publisher>Virginia Polytechnic Institute and State University</publisher>
<date>1997-03-31</date>
<type>text</type>
<format>application/pdf</format>
<identifier>http://scholar.lib.vt.edu/theses/available/etd-3345131939761081/</identifier>
<language>en-us</language>
<rights>unrestricted</rights>
</oai_dc>
</metadata>
</GetRecord>
Open Archives Initiative Repository Explorer
v1.0

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

Archive details: http://etd.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/vtetd/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
• Distributed DL for educational material

• CSCT and I Lumina

• Central digital document archive

• VT Digital Library

ongoing projects 2
Vision

Lots and lots of little archives all over the world

OAI

International ETD Archive

International Teaching Resources Archive

Etc ...