Customized Mapping and Metadata Transfer from DSpace/SOAR to OCLC to Improve ETD Work Flow

- Sai Deng, Susan Matveyeva, Tse-Min Wang, Wichita State University Libraries
- Consultant: Terry Reese, Oregon State University Libraries
Outlines

- Thesis Cataloging Workflow Dynamics: overview of changes
- Cataloging ETDs in SOAR and OCLC/Voyager: records & workflow
- Improving ETD Workflow through metadata harvesting, customized mapping and metadata transfer
Workflow for Paper Theses

- 1929-2002 – over 80% records (~ 5000)
- 70 year range: stable record’s structure
- Workflow: (1) original cataloging (2) item’s marking/labeling
- Cataloging efficiency: constant data
- Labor intensive: SH
A descriptive study of obstetric patients' knowledge of and self reported attitudes toward the prenatal experience / by Martha Tasheff Vliet.


viii, 75 leaves ; 29 cm.

Also in University Archives: THESIS.

Title on spine: Patients’ perceptions of prenatal experience.


Bibliography: leaves 48-52.

Wichita State University.

Pregnancy.

Pregnancy Psychological aspects.

Prenatal care.
2003-2004 digitization of WSU Theses began

UMI/ProQuest effects workflow

Linking Voyager records to UMI/ProQuest
Record enhancements (fields /contents)

856 - links from a catalog to full text in UMI

520 – author abstracts

500 & 700 -- advisor’s name

Workflow changes: Special projects: a repetitive data entry goes to students

Cataloger creates procedure; MACRO for speedy processing; trains students, and review their work
A comparison of multiple-stage tandem MS of protonated and metal cationized peptides in the context of direct sequencing and sequence tag generation / by Bupani Asiri Perera.

Thesis advisor: Michael J. Vanstipdonk.

Includes bibliographical references (leaves 128-136).

We have examined the multiple stage collision we bind to the metal ion significantly.

Click here for available full-text of this dissertation via Current Research@Gateway.
Transitional Period: 2004-2006

- e-Theses in four places: OCLC/Voyager; ProQuest; a temporary web site and SOAR
- Paper theses are still submitted
- Development of a new workflow for ETDs
- e-docs, paper docs, inventory table
- Naming convention, ETD file preparation
- MARC and DC manual input; further changes in records (identifiers)
ETD Program 2006-2008

- From 2006, WSU have a full scale ETD program (400 records, 2005-2007)
- eTheses (no paper); no ProQuest or temporary access to ETD via a web site
- eTheses are in three databases: SOAR and OCLC/Voyager
- Work Flow includes the number of operations with a digital file (thesis) and metadata records (MARC and DC)
<table>
<thead>
<tr>
<th>Pdf ID</th>
<th>No</th>
<th>Last First Name</th>
<th>Year</th>
<th>Mon.</th>
<th>GS send list</th>
<th>PDF Harvested</th>
<th>PDF Property filled</th>
<th>PDF Subm To UMI</th>
<th>PDF secured</th>
</tr>
</thead>
<tbody>
<tr>
<td>d07001</td>
<td>1</td>
<td>Smith John</td>
<td>2007</td>
<td>May</td>
<td>date</td>
<td>date</td>
<td>date</td>
<td>date</td>
<td>date</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PDF re-named</th>
<th>GS Paper work received</th>
<th>Soar ID</th>
<th>Voyager Bib</th>
<th>UMI ID</th>
<th>UMI Link</th>
<th>Soar Link</th>
<th>Micr film No</th>
<th>Link Checked</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>date</td>
<td>date</td>
<td>1074</td>
<td>1262388</td>
<td>3240865</td>
<td>Yes/no</td>
<td>Yes/no</td>
<td>2740</td>
<td>date</td>
<td></td>
</tr>
</tbody>
</table>
ETD Workflow: Manual Input DC & MARC

1. Create e-folders for PhD/MS
2. *Create inventory table in MS Excel
3. Download from Blackboard
4. Paper work
5. Create e-folders for PhD/MS

- create printouts of title pages
- alphabetize place in binder
- Label title printouts with file name

- Send through FTP to UMI/ProQuest
- Rename ETD pdfs
- Fill out properties
- Secure with password

- Create DC in SOAR
- Create MARC in OCLC

- Upload to SOAR
- Link to ProQuest
- Link to SOAR
- Link to ProQuest
- Import MARC to Voyager

NOTE: Each step of workflow reflected in inventory table
The Improved Workflow: no draft record and manual MARC input
A Wider Context of ETD Workflow

ETD workflow in different institutions
  - Home-grown scripts, site-specific harvesters
- Kent State University (2007)
  - Harvest from OhioLINK ETD Center, ETD-MS to Marc...

XSLT Transformation
- LC MARC 21 XML schema with MarcXML toolkit
  - Dublin Core to MARCXML Stylesheet
- OAI community developed tools, mostly for IT staff
  - MarcEdit (Terry Reese)
    - Metadata Harvester, MARC Editor
    - Low-barrier harvester, can be used by catalogers
<table>
<thead>
<tr>
<th>DC Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dc.contributor.author</td>
<td>Niles, Rae-</td>
</tr>
<tr>
<td>dc.date.accessioned</td>
<td>2006-12-24T14:56:10Z</td>
</tr>
<tr>
<td>dc.date.available</td>
<td>2006-12-24T14:56:10Z</td>
</tr>
<tr>
<td>dc.date.copyright</td>
<td>2006</td>
</tr>
<tr>
<td>dc.date.issued</td>
<td>2006-05</td>
</tr>
<tr>
<td>dc.identifier.other</td>
<td>d06005</td>
</tr>
<tr>
<td>dc.identifier.uri</td>
<td><a href="http://hdl.handle.net/10057/373-">http://hdl.handle.net/10057/373-</a></td>
</tr>
<tr>
<td>dc.description</td>
<td>Thesis (Ed.D.)--Wichita State University, College of Education.en</td>
</tr>
<tr>
<td>dc.description</td>
<td>&quot;May 2006.&quot;</td>
</tr>
<tr>
<td>dc.description.abstract</td>
<td>The purpose of this study was to describe and identify Sedgwick High School’s teacher and student perceptions of the impact of one-to-one laptop computer access using an appreciative inquiry theoretical research perspective and the theoretical frameworks of change and paradigm shift...</td>
</tr>
<tr>
<td>dc.format.extent</td>
<td>1174852 bytes-</td>
</tr>
<tr>
<td>dc.format.mimetype</td>
<td>application/pdf-</td>
</tr>
<tr>
<td>dc.language.iso</td>
<td>en_US</td>
</tr>
<tr>
<td>dc.rights</td>
<td>Copyright Rae Niles, 2006. All rights reserved.</td>
</tr>
<tr>
<td>dc.subject.lcsh</td>
<td>Educational technology</td>
</tr>
<tr>
<td>dc.subject.lcsh</td>
<td>Education--Data processing</td>
</tr>
<tr>
<td>dc.subject.lcsh</td>
<td>Electronic dissertations</td>
</tr>
<tr>
<td>dc.title</td>
<td>A study of the application of emerging technology: teacher and student perceptions of the impact of one-to-one laptop computer access</td>
</tr>
<tr>
<td>dc.type</td>
<td>Dissertation</td>
</tr>
<tr>
<td>dc.thesis.adviser</td>
<td>Calabrese, Raymond L.</td>
</tr>
<tr>
<td>dc.identifier.oclc</td>
<td>71805797-</td>
</tr>
<tr>
<td>Appears in Collections:</td>
<td>EL Theses and Dissertations</td>
</tr>
<tr>
<td></td>
<td>COE Theses and Dissertations</td>
</tr>
<tr>
<td></td>
<td>Dissertations</td>
</tr>
</tbody>
</table>
Dublin Core to MARC Mapping

Fields in DSpace                  Transformed MARC fields in OCLC (What we want)

- dc.contributor.author  ➔  100 1 _  Author.
- dc.date.accessioned ➔
- dc.date.available ➔
- dc.date.copyright ➔
- dc.date.issued ➔
- dc.identifier.other ➔
- dc.identifier.uri ➔
- dc.description ➔  260  ≠c year.
- dc.description ➔  099  ……
- dc.description ➔  856 4 0 …
- dc.description ➔
- dc.description.abstract ➔
- dc.format.extent ➔
- dc.format.extent ➔
- dc.format.mimetype ➔
- dc.language.iso ➔
- dc.rights ➔
- dc.subject ➔
- dc.subject.lcsh ➔
- dc.title ➔
- dc.type ➔
- dc.thesis.adviser ➔
- dc.identifier.oclc ➔
- Appears in Collections: ➔

546  en_US
540  Access restricted to WSU students, faculty and staff (delete)
690  (keywords, non CV, delete)
650  _ 0
245 1 _  …
655 7  Dissertation ≠2 local
700 1 2 … ≠e advisor
856 4 1 …
Using MarcEdit

MarcEdit Interface

Metadata Harvester

Harvest OAI Data

Server Address: http://soar.wichita.edu:8080/oispace-oai/request

Set Name:

Metadata Type: Dublin Core

Crosswalk Path: C:\Program Files\MarcEdit5.0\ws\mr\OAIDToMARCXML.xsl

Advanced Settings

OK  Close
Metadata transformation in MarcEdit

- The wheel and spoke design for metadata transformation (by Reese)

Diagram:
- Dublin Core
- EAD
- TEI
- MODS

Interconnections:
- Dublin Core ↔ EAD ↔ MODS ↔ TEI
Data Flow Diagram

- **DSpace**: OAI request → OAI response
- **OCLC**
- **MarcEdit**
  - Metadata Harvester
  - MarcEditor
  - **XSLT** (DC to MarcXML)
- **Voyager**: Export MARC

**Authorized data processing**
(Title, author, subject…)

**Resolving data ambiguity**
(Many to one mapping w/ element positioning…)

**String Processing**
(Data normalization…)
Selective Harvesting

- Define in MarcEdit
  - by identifier (e.g. oai:soar.wichita.edu:10057/255)
  - by set (e.g. hdl_10057_351)
  - by date (e.g. from=2007-01-01&until=2008-01-01)


- How do we define harvesting theses only?
  - Define by set (http://soar.wichita.edu/dspace-oai/request?verb=ListSets)
    - Sets by schools and departments
      - AE Theses and Dissertations (hdl_10057_313)
      - ANTH Theses (hdl_10057_233)
      - BIO Theses (hdl_10057_389)
      - CE Theses and Dissertations
      - ...
    - Or sets in two categories
      - Master's These (hdl_10057_351)
      - Dissertations (hdl_10057_352)
Alternatively, Define Theses Sets in XSLT

**Dublin Core to MARCXML Stylesheet**

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet version="1.0" xmlns:dc="http://purl.org/dc/elements/1.1/
xmlns:dcterms="http://purl.org/dc/terms/1.1"
xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd" xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
xmlns="http://www.loc.gov/MARC21/slim" exclude-result-prefixes="dc dcterms oai_dc">
<xsl:import href="MARC21slimUtils.xsl" />
<xsl:output method="xml" encoding="UTF-8" indent="yes" />
- <xsl:template match="/">
  <collection xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/MARC21/sl
http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd">
    <xsl:apply-templates />
  </collection>
</xsl:template>
- <xsl:template name="OAI-PMH">
  - <xsl:for-each select="ListRecords/record/metadata/oai_dc:dc">
    <xsl:apply-templates />
  </xsl:for-each>
</xsl:template>
- <xsl:for-each select="GetRecord/record/metadata/oai_dc:dc">
  <xsl:apply-templates />
</xsl:for-each>
</xsl:template>
```
XSLT Customization: Transform and Display Theses and Dissertations Only

<record>
  <xsl:variable name="myType" select="dc:type" />
  - <xsl:choose>
    - <xsl:when test="($myType!="") and (($myType='Thesis') or ($myType='Dissertation'))">
      - <xsl:element name="leader">
        <xsl:variable name="type" select="dc:type" />
        - <xsl:variable name="leader06" />
      </xsl:element>
    </xsl:when>
  </xsl:choose>
  <xsl:when test="$type='collection'">p</xsl:when>
  <xsl:when test="$type='dataset'">m</xsl:when>
  <xsl:when test="$type='event'">r</xsl:when>
  <xsl:when test="$type='image'">k</xsl:when>
  <xsl:when test="$type='interactive resource'">m</xsl:when>
  <xsl:when test="$type='software'">m</xsl:when>
  <xsl:when test="$type='service'">m</xsl:when>
  <xsl:when test="$type='sound'">i</xsl:when>
  <xsl:when test="$type='text'">a</xsl:when>
  <xsl:when test="($type='Thesis') or ($type='Dissertation')">t</xsl:when>
  <xsl:otherwise>a</xsl:otherwise>
</xsl:choose>
...
A study of the application of emerging technology, teacher and student perceptions of the impact of one-to-one laptop computer access

The purpose of this study was to describe and identify Sedgwick High School's teacher and student perceptions of the impact of one-to-one laptop computer access using an appreciative inquiry theoretical research perspective and the theoretical frameworks of change and paradigm shift. An appreciative inquiry theoretical research perspective was used to structure a qualitative, embedded descriptive case study design. An embedded case study design was used to describe the perceptions of high school teachers and their students who were involved in a one-to-one laptop computer wireless environment on student learning and how teachers teach. Data were collected through teacher and student focus groups, as well as administration of the Left Hand Right Hand Column Case Method. Data were analyzed using the Comparative Analysis Matrix method (Miles & Huberman, 1994). The analysis data revealed six salient findings: (1) Students functioned in the capacity of teacher, (2) technology changed the way teachers and students communicated, (3) the culture of the classroom dynamics between teacher and student changed, (4) technology made learning enjoyable for students, (5) teachers and students believed immersion in a technology-rich learning environment created advantages for student success after high school graduation, and (6) teachers believed that access to ubiquitous technology created new challenges for maintaining student engagement in the learning process. Five of the six findings suggested that technology had changed teaching and learning, and helped to create a paradigm shift in the teacher and student roles. Additionally, those findings also identified the positive role of Sedgwick High School, serving to describe the little giving forces within the organization. One of the six findings revealed challenges associated with the application of emerging technology in the classroom. The findings from this study have the potential to contribute to areas of study that focus on the use of technology in schools. Moreover, research from this study has the potential to help serve as a foundation for other school leaders who are seeking opportunities that prepare students for life in the technology-rich 21st Century through one-to-one laptop computer access.
Mapping Problems and Error Reports (for Variable Fields)

- 100 occurrence 1, indicator 2 - invalid code
- 520 occurrence 4, $a occurrence 1, position 76 - invalid character - data must be ALA characters
- 655 occurrence 1, indicator 1 - invalid code
- 655 occurrence 1, indicator 2 - invalid code
- 655 occurrence 1, $2 - invalid relationship - when element is present, then 655 indicator 2 must equal 7

...

Need customization to meet our needs.
DSpace (version 1.4 or below) only responds with simple Dublin Core xml file (to be transformed to MarcXML using xslt).

Fields in DSpace  Transformed fields in OCLC  Correction and Customization Needed
---  ---  ---
dc.contributor.author  100 1 0 Niles, Rae  +e author  (Delete +e author.)
dc.date.accessioned  

dc.date.available  

dc.date.copyright  

dc.date.issued  260  +c 2006-05  (Only keep 2006)
dc.identifier.other  500  d06005  (Change to 099)
dc.identifier.uri  500  http://hdl.handle.net/10057/373  (Change to 856 4 0)
dc.description  520  Thesis (Ed.D.)--Wichita State University, College of Education.  (Change to 502)
dc.description  520  "May 2006."  (Change to 500)
dc.description  520  Includes bibliographic references (leaves 129-145).  (Change to 504)
dc.description.abstract  520  The purpose of this study was to describe and identify Sedgwick High School’s teacher and student perceptions of the impact of one-to-one laptop computer access using an appreciative inquiry theoretical research perspective and the theoretical frameworks of change and paradigm shift...  (Change to 520 3)
dc.format.extent  

dc.format.extent  

dc.format.mimetype  

dc.language.iso  546  en_US  (delete)
dc.rights  540  Access restricted to WSU students, faculty and staff  (delete)
dc.subject.lcsh  690  Educational technology  (Change to 650 _0)
dc.subject.lcsh  690  Education--Data processing

dc.subject.lcsh  690  Electronic dissertations

dc.title  245 0 0 A study of the application of emerging technology: teacher and student perceptions of the impact of one-to-one laptop computer access  (if 100 exists, use 245 1_; or else use 245 0_)
dc.type  655 7  Dissertation +2 local  (Change to 655 _7)
dc.thesis.adviser  

dc.identifier.oclc  856 4 1  +u 71805797  +z Connect to this object online.  (replace +u with value from dc.identifier.uri)

Appears in Collections:  

Resolving data ambiguity

- Same DC fields to different MARC fields:
  - description → 502 (Dissertation)
    500 (General Note)
    504 (Bibliography)

- Qualified DC element:
  - description.abstract → 520 (Summary)

Solution: element positioning

```xml
<xsl:for-each select="dc:description[1]">
    - <datafield tag="502" ind1="" ind2="">
        - <subfield code="a">
            <xsl:value-of select="normalize-space(.)" />
        </subfield>
    </datafield>
</xsl:for-each>

<xsl:for-each select="dc:description[2]">
    - <datafield tag="500" ind1="" ind2="">
        - <subfield code="a">
            <xsl:value-of select="normalize-space(.)" />
        </subfield>
    </datafield>
</xsl:for-each>
```

...
Authorized data processing

- Primary entries vs. added entries: title and personal names processing
  - Template to deal with personal names (in MarcEdit)
    - E.g. `<dc:creator>Webb, Kyle M.</dc:creator>`
      transformed to:
      - `<dc:creator>Webb, Kyle M.</dc:creator>`

- Identify field relationship and correct indicators
  - 100, 245 (author, title) relationship: if 100 exists, 245 1 _
    or else, 245 0 _

- Local element: `dc.thesis.advisor` transformed to 700 1_
  (If more than one dc.thesis exists, positioning is needed.)
Customized Mapping in XSLT

- Processing of non-filing characters in title
  - 245 (title) 2nd indicator: a, an, the… (2, 3, 4)
    <xsl:for-each select="dc:title[1]">
      - <xsl:choose>
        - <xsl:when test="$exist100!=''">
          - <xsl:choose>
            - <xsl:when test="substring(., 1, 2)='A '">
              - <datafield tag="245" ind1="1" ind2="2">
                - <xsl:choose>
                  - <xsl:when test="contains(.,':')">
                    - <subfield code="a">
                      <xsl:value-of select="concat(substring-before(.,':'),' : ')
                    </subfield>
                    - <subfield code="b">
                      <xsl:value-of select="concat(substring-after(.,':'),' / ')
                    </subfield>
                  </xsl:when>
                  …
                </xsl:choose>
              </datafield>
            </xsl:when>
            <xsl:when test="contains(.,':')">
              <subfield code="c">
                <xsl:value-of select="concat(substring-before(.,':'),' : ')
              </subfield>
              <subfield code="d">
                <xsl:value-of select="concat(substring-after(.,':'),' / ')
              </subfield>
            </xsl:when>
          </xsl:choose>
        </xsl:when>
        <xsl:when test="contains(.,';')">
          <xsl:value-of select="concat(substring-before(.,';'),', ')
        </xsl:when>
      </xsl:choose>
    </xsl:for-each>

- Alternatively, it can be defined in the title template.
Customized Mapping in XSLT

- **Subjects vs. Keywords**
  - Only kept common subject in the test (when keywords and subjects mixed inconsistently)
    - `<xsl:for-each select="dc:subject">
        - `<xsl:if test="."="Electronic dissertations">
            - `<datafield tag="650" ind1="" ind2="0">
                - `<subfield code="a">
                    - `<xsl:value-of select="." />`
                </subfield>
            </datafield>
        </xsl:if>`
    </xsl:for-each>`

- **Subject template** (OSU solution)
  - `<dc:subject>ocean wave energy</dc:subject>`
  - `<dc:subject>direct-drive</dc:subject>`
  - `<dc:subject>fluid-structure interaction</dc:subject>`
  - `<dc:subject>Ocean wave power</dc:subject>`
  - `<dc:subject>Fluid-structure interaction</dc:subject>`

  Transformed to
  
  - 650 \$aOcean wave power.
  - 650 \$aFluid-structure interaction.
  - 690 \$aelectric wave energy.
  - 690 \$adirect-drive.
  - 690 \$afuid-structure interaction.
Customized Mapping in XSLT

String Processing

Functions

- normalize-space()
- translate()
- substring()…

Example: Extract partial value from DC element

- 260 (Date): only extract year from the issuing date in DC

```xml
  - <xsl:for-each select="dc:date[4]">
    - <xsl:if test=".!=''">
      - <datafield tag="260" ind1='' ind2=''">
        - <subfield code="c">
          <xsl:value-of select="substring(.,1,4)" />
        </subfield>
      </datafield>
    </xsl:if>
  </xsl:for-each>
```
Customized Mapping in XSLT

- **Leaders:** fixed fields that comprise the first 24 character positions (00-23) of each MARC record. They provide information for the processing of the record.

- **008 field (Fixed-Length Data Elements)**
  - Type (t, manuscript language material)  
  - Desc (a)  
  - Cont (b, m, content is theses with bibliographies)  
  - Srce (d, cataloging source)  
  - Fest (0, not a festschrift)  
  - Indx (0, no index)  
  - BLvl (m, Encoding level is monograph)  
  - ELvl (I, encoding level is full level)  
  - Form (s, form of item is electronic)  
  - Ills (a, illustration included)  
  - Conf (0, not a conference publication)  
  - LitF (0, not fiction)  
  - DtSt (s, single date)  
  - Lang (eng, language is English)  
  - Ctry (xx)

- **Ways to handle:**
  - Scripting and adding all fixed fields (leader and 008 fields) in OAIDCtoMARCXML.xsl;
  - Or, Adding 008 in MarcEditor after record export;
  - Or, applying fixed field template after records being exported to OCLC.
Harvesting Using the Revised XSLT Crosswalk

Harvest Raw Data

Metadata Harvester

- Server Address: http://soar.wichita.edu:8080/dspace-oai/request
- Set Name: 
- Metadata Type: Dublin Core
- Save Folder: C:\Documents and Settings\saineng\Desktop\MarcEditchеr

Advanced Settings

- GetRecord:
- ResumptionToken:
- Start: 
- End: 
- Timeout: 100 secs.
- Harvest Raw Data (save OAI data to local file system)

OK  Close
Raw DC XML (Harvest oai Data to Local File)

<?xml version="1.0" encoding="UTF-8"?>
<record>
  <header>
    <identifier>oai:soor.wichita.edu:10057/264</identifier>
    <datestamp>2008-03-04T14:37:00Z</datestamp>
  </header>
  <metadata>
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc.dtd">
      <dc:creator>Malarkar, Bijaya</dc:creator>
      <dc:date>2006-11-23T04:20:58Z</dc:date>
      <dc:date>2006-11-23T04:20:58Z</dc:date>
      <dc:date>2006-11-23T04:20:58Z</dc:date>
      <dc:date>2006-11-23T04:20:58Z</dc:date>
      <dc:date>2006-11-23T04:20:58Z</dc:date>
      <dc:identifier>http://hdl.handle.net/10057/264</dc:identifier>
      <dc:creator>Thesis (M.S.)--Wichita State University, Dept. of Mechanical Engineering.</dc:creator>
      <dc:description>Includes bibliographic references (leaves 73-75).</dc:description>
      <dc:description>Bipedal walking robots have distinguished themselves from other robots not only for their better mobility but also for the increased amount of complexity in their architecture. The issue of walking stability has inherently been related with these in various aspects from dynamics to controls. It has been defined in various ways for different aspects and these definitions have served as guidelines for designing a walking robot. This research is based on a design of a walking robot and development of a walking trajectory. Among the various approaches implemented in order to attain control of two-legged walking, this study attempted to linearize the walking trajectory by dividing it into finite discrete sections. Using the method of kinematics inversion to generate the initial open loop path, the study looks into implementation of the results for a successful walking mechanism. Further it dealt with stepping velocity control by using ground contact trigger and aimed at reducing the stress developed in joints due to dynamic forces.</dc:description>
      <dc:format>pdf</dc:format>
      <dc:rights>Copyright Bijaya Malarkar, 2006. All rights reserved.</dc:rights>
      <dc:subject>Electronic dissertations</dc:subject>
    </oai_dc:dc>
  </metadata>
</record>
Harvest and Transform DC to MarcXML

Metadata Harvester

Harvest OAI Data
- Server Address: http://soar.wichita.edu:8080/dspace-oai/request
- Set Name: 
- Metadata Type: Dublin Core
- Crosswalk Path: C:\Program Files\MarcEdit 5.0\XSLT\OAIDtoMARCXML.xsl

Results
- Total Records Processed: 1000
- ResumptionToken Processed: 10
- Last ResumptionToken Processed: 0001-01-01T00:00:00Z/9699-12-31T23:59:59Z/oai_dc/1000

Advanced Settings

OK  Close
Records will be Dumped to MarcEdit-MarcEditor
MarcEditor

- **Edit harvested theses in MarcEditor**
  - Batch edit fields, subfields, indicators (if needed)
  - E.g.: add 008 field for all records
- .mrk (MARC text file) \(\rightarrow\) Compile to .mrc (MARC)

Or

- Save as .mrk8 (MARC UTF8 text file) \(\rightarrow\) Compile to .mrc (MARC)
Import Records to OCLC

- Click “File-Import Records…”
- Select “Import to Local Save File”
<table>
<thead>
<tr>
<th>Record ID</th>
<th>Contribution/Creator</th>
<th>Title</th>
<th>Control #</th>
<th>Date</th>
<th>Call Number</th>
<th>Date/Time Added</th>
<th>Held Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Golbein, Albert</td>
<td>Title?</td>
<td>NE0001</td>
<td>6/22/2008 10:26:51 AM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Shahidi, Audrey Dawn author</td>
<td>The effect of line length and passage type on reading.</td>
<td>NE0002</td>
<td>6/22/2008 2:50:04 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Singamajjula, Sushadrakumar K.</td>
<td>Signal processing of acoustic reflection in estima...</td>
<td>NE0003</td>
<td>6/22/2008 2:50:04 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Potter, unknown author</td>
<td>San Francisco bowl</td>
<td>NE0004</td>
<td>6/22/2008 2:50:05 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Huest, Brian Daniel</td>
<td>Where did I park? Connecting lower-level and higher-...</td>
<td>NE0005</td>
<td>6/3/2008 3:5:13 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Millen, Fae</td>
<td>Study of the application of emerging technology: team...</td>
<td>NE0006</td>
<td>6/3/2008 3:5:13 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Smith, Phillip M.</td>
<td>Studies on electron and energy transfer in porphyrin...</td>
<td>NE0007</td>
<td>6/3/2008 3:5:13 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Miller, Russell K.</td>
<td>Impact of educational technology on learner interaction...</td>
<td>NE0008</td>
<td>6/3/2008 3:5:13 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Gabai, Suresh</td>
<td>Supramolecular poly(propylene carbonate) design...</td>
<td>NE0009</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Mosey, Victoria</td>
<td>Student ratings of university teaching: similarities and...</td>
<td>NE0010</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Kim, Young-Min</td>
<td>Robust and reduced order H-Infinity filtering via LMI ap...</td>
<td>NE0011</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Hyun, Inha</td>
<td>Stochastic control of unlined decentralized singularly p...</td>
<td>NE0012</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Kumbhar, Sachin Shivaji</td>
<td>Development of a finite element model and analysis of...</td>
<td>NE0013</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Summers, Bryce</td>
<td>Effects of family structure and parenting style on cho...</td>
<td>NE0014</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Swartzendruber, Rachel D.</td>
<td>Discovering voices among peculiar quietness: an aud...</td>
<td>NE0015</td>
<td>6/3/2008 3:5:11 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Binns, Rebecca Kay</td>
<td>&quot;On the cover of a Rolling Stone&quot;: a content analysis o...</td>
<td>NE0016</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Origum, Recep</td>
<td>Design and timing analysis of wave pipelined circuits...</td>
<td>NE0017</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Ott, Sara</td>
<td>Paradox and philosophical anticipation in Melville's M...</td>
<td>NE0019</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Shere, Sudhir Shankar</td>
<td>Energy absorption of a car roof reinforced with a grid...</td>
<td>NE0020</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Goenka, Supriya Shanika</td>
<td>Effect of large neutral amino acids on maternal phenyl...</td>
<td>NE0021</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Davis, Thomas Arthur</td>
<td>Evolution of literary theory: towards a bi-cultural appr...</td>
<td>NE0022</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Shiflett, Thomas Robert</td>
<td>West Nile virus and wild bird populations</td>
<td>NE0023</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Elsayed, Tazrane</td>
<td>United procedure for continuous-time and discrete-time...</td>
<td>NE0024</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Dajihmazian, Bilyar</td>
<td>Effect of variations in riveting process on the quality...</td>
<td>NE0025</td>
<td>6/3/2008 3:5:12 PM</td>
<td>KSMA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
After Being Exported to OCLC…

In OCLC Connexion client:
Open each file, do some review/editing as needed, attach KSW holding and apply fixed field template of ETD (if needed) in OCLC.
Alternatively, records exported to Voyager directly

- This part is performed by Gemma Blackburn.
- Send .mrc file to the Voyager server.
- Create a Bulk Import rule in Voyager System Administration module.
  - Go to: Cataloging → Bulk Import Rules → New
  - Name the rule
  - Choose (or create a new) Bib De-Duplication Rule
  - Modify mapping as needed
  - Save the rule
Voyager System Administration Bulk import rules screenshot
Bulk Import the records using the Bulk Import rule

- On your Voyager server, go to: .../voyager/xxxdb/sbin/
- Write the command for Bulk Import to run: Pbulkimport –ftheses-sample.mrc –iSOAR –b1 –e3
  - –f and the file name (required)
  - –i and the Bulk Import rule name (required)
  - –o and your name (not required, but will let people know who ran the bulk import)
  - –b and a number. This will define the beginning record in the file that you want to import if you prefer to import a select set at a time (not required)
  - –e and a number. This will define the end record in a set to import (not required)
- There are several other options. Check the Technical User’s Guide
A real case

- Transformation of ETDs of 2007
  - Ph.D. Dissertations (Summer, Fall 2007): 23
  - Master’s Theses (Summer, Fall 2007): 55

- Some adjustment in the transformation:
  - Transfer dc.format.extent[1] to physical description (Marc 300)
    - E.g. ix, 53 leaves, ill. → 300 $a ix, 53 leaves : $b ill.

- Keep 3 description fields
  - description [1] → 500(General Note)
  - description.abstract → 520(Summary)

- 008 field values added in MarcEditor rather than applied in OCLC
  - E.g. =008 ...s2008\\xx\\\\sbm\\000\0\eng\d
Discussion and Conclusion

- The customized mapping and metadata transfer can eliminate the need of double entry in DSpace and OCLC/Voyager and significantly improve our ETD work flow.

- Metadata management
  - One single crosswalk and style sheet will not meet all needs;
  - Needs to be based on standard practice but add local variations;
  - Application-specific mapping is needed for special projects;
  - Coordination in metadata repurposing is important.

- Data mapping, manipulation and transformation
  - Using qualified DC instead of element positioning in XSLT;
  - DSpace 1.5 enables qualified DC crosswalk for OAI-PMH;
  - Handling of MARC fixed fields and 008 field.

- Other technical issues
  - Using other tools for harvesting besides MarcEdit;
  - Using DSpace Item Importer and Exporter instead of Metadata Harvester.
Project team and Acknowledgement

- Sai Deng, Metadata mapping and transformation
- Susan Matveyeva, ETD cataloging and mapping
- Tse-Min Wang, Programming assistance
- Sandy Oswald, Manoj Gogoi, ETD cataloging assistance
- Terry Reese, Consultant
- Nancy Deyoe, Administrative Support
- Connie, Basquez, Voyager support
- Gemma Blackburn, Voyager support
Thank you!