Building Virtual Collections

by Nancy Stuart
University of Victoria

For the BCLA pre-conference
Beyond Limits: Building Open Access Collections
April 19, 2007
Introduction

- This session will examine the underpinnings of building virtual collections; metadata and application profiles
- Who I am – Technical Services Librarian and ETD Project Manager at University of Victoria and member of the CARL IR Metadata Working Group
Session Outline

- CARL vision
- CARL harvester
- Metadata
- Application profiles
- CARLCore
- Institutional Repositories
- Challenges
CARL vision for virtual collections

- CARL Institutional Repository Project
- Federated searching
- Scholarly publishing
- Open Access environment
Virtual collections

- Definition
- Small or large
- Text, images, maps
- Big picture
# Print to Digital

**Print**

<table>
<thead>
<tr>
<th>Description</th>
<th>Local catalogue</th>
<th>Union catalogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Points AACR2/MARC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Books – stacks</td>
<td>author</td>
<td>card</td>
</tr>
<tr>
<td>Maps – drawers</td>
<td>title</td>
<td>OPAC</td>
</tr>
<tr>
<td>Videos – film desk</td>
<td>subject</td>
<td></td>
</tr>
</tbody>
</table>

**Digital**

<table>
<thead>
<tr>
<th>Description</th>
<th>Local catalogue</th>
<th>Union catalogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata DC/VRA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Books – PDFs</td>
<td>author</td>
<td>OPAC</td>
</tr>
<tr>
<td>Images - JPGs</td>
<td>title</td>
<td>DSpace</td>
</tr>
<tr>
<td>Music – MP3s</td>
<td>keyword</td>
<td>local DB</td>
</tr>
<tr>
<td>Best practices</td>
<td>AP</td>
<td></td>
</tr>
</tbody>
</table>

Repositories & Harvesters

- What is a digital repository?
  - Data provider
  - Contains electronic texts, maps, learning objects, etc. & metadata

- What is a harvester?
  - Service provider
  - Harvests the metadata
  - Single search interface
CARL harvester

- Established Mar. 2004
- Hosted at Simon Fraser University
- Uses OAI-PMH
- Currently 12 institutions – 12 more coming online this summer
- April 2007, over 30,000 records
Metadata

- MARC is a metadata schema; coded and complex
- Dublin Core is a metadata schema designed for digital resources; plain language and simple
- Consistency in values, such as Date and Type, affects resource discovery
Application profiles

- A defined element set, policies and “best practices”
- Standard vocabularies, eg. DCMI Type vocabulary
- Guidelines
- CARLCore - CARL harvester application profile
<table>
<thead>
<tr>
<th>Name of Term</th>
<th>dateCopyrighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term URI</td>
<td><a href="http://purl.org/dc/terms/dateCopyrighted">http://purl.org/dc/terms/dateCopyrighted</a></td>
</tr>
<tr>
<td>Label</td>
<td>Date</td>
</tr>
<tr>
<td>Defined By</td>
<td><a href="http://dublincore.org/documents/dcmi-terms/">http://dublincore.org/documents/dcmi-terms/</a></td>
</tr>
<tr>
<td>Source Definition</td>
<td>Date of a statement of copyright</td>
</tr>
<tr>
<td>Source Comments</td>
<td></td>
</tr>
<tr>
<td>[Domain] Comments</td>
<td>Use the year of copyright from the thesis title page.</td>
</tr>
<tr>
<td>Type of term</td>
<td>element-refinement</td>
</tr>
<tr>
<td>Refines</td>
<td>date</td>
</tr>
<tr>
<td>Refined By</td>
<td>M</td>
</tr>
<tr>
<td>Obligation</td>
<td></td>
</tr>
<tr>
<td>Occurrence</td>
<td>___</td>
</tr>
</tbody>
</table>
CARLCore

- CARLCore Metadata Application Profile
- Collaborative effort of members across Canada
- Based on unqualified Dublin Core
- Version 1.0 released mid May
Institutional Repositories

- Create their own metadata application profiles
- Summarize any follow up action items required of you
- CARL encourages members to implement IRs
- CARL harvester is one common point of access to Canadian IRs
CARLCore element - Description

Element: Description
Definition: An account of the content of the resource.
Obligation: Optional
Generated By: Person
Recommended Encoding: None
Element Guidelines: Repeatable.
      Free text.
      May include an abstract, table of contents, summary description of the contents, or a URI which points to a description.

Examples:
Illustrates the use of a Webcam. [Image]
This thesis describes the high velocity vibration of the proton ... [Text]
Challenges to building virtual collections

- Canadian context means French and English
- Author generated metadata
- Collective vs. institutional level
- Diverse and decentralized
More information

- Nancy Stuart: nstuart@uvic.ca
- CARL harvester at: http://carl-abrc-oai.lib.sfu.ca/
- CARL website: http://www.carl-abrc.ca
- CARL IR Project http://www.carl-abrc.ca/projects/institutional_repositories/institutional_repositories-e.html