Research Data Management
At the University of Pretoria

ID van der Walt – July 2014
Overview

1. ECM Approach

2. Supporting the Research Cycle

3. Alfresco
   • Live Demo

1. Bagit
“Enterprise Content Management (ECM) is a formalized means of organizing and storing an organization's documents, and other content, that relate to the organization's processes. The term encompasses strategies, methods, and tools used throughout the lifecycle of the content.”


**To Note:**

- Enterprise content management is not a closed-system solution or a distinct product category.
- Focus should be focused on your environment.
Supporting the Research Cycle

- Research Planning and Design
- Data Collection – Short term storage and sharing
- Data Study and Analysis
- Data Publishing
- Long Term Storage and Preservation
- Reuse of Data
Supporting the Research Cycle Cont.
Supporting the Research Cycle Cont.

1. Research In Action – 75%
   • Alfresco

2. Dissemination
   • Dspace

3. Preservation
   • Bagit
Alfresco – Live Demo
What is BagIt

1. “BagIt is a hierarchical file packaging format designed to support disk-based storage and network transfer of arbitrary digital content.“

1. “A "bag" consists of a "payload" (the arbitrary content) and "tags", which are metadata files intended to document the storage and transfer of the bag.“

1. “They are also well-suited to the export, for archival purposes, of content normally kept in database structures that receiving parties are unlikely to support.“

http://en.wikipedia.org/wiki/BagIt
Contents of a Bag

1. Data Directory
   • Contains the data payload
   • Can be single or multiple files and directories

1. Manifest File
   • Text file with listed items in the payload with their checksums

1. BagIt File
   • Contains info on BagIt version and encoding

2. BagIt Info File
   • Contains the Metadata for the bag

3. Tag Manifest File
   • Contains checksums to verify the above mentioned txt files
Example

About Bagger 2.1.2

Bagger 2.1.2

Description
The Bagger application creates and deposits bags using the BagIt specification.

Bagger Application
Version: 2.1.2

Bag It Library
Version: 3.13

Distribution Site
http://sourceforge.net/projects/loc-xferutils/
Example
Example

Save Bag Dialog

Save Bag
Define the Bag settings

Save in:

Holey Bag?:

Base URL

Serialize Type?:

Generate Tag Manifest?:

Tag Manifest Algorithm:

Generate Payload Manifest?:

Payload Manifest Algorithm:

OK Cancel
<table>
<thead>
<tr>
<th>Name</th>
<th>Date modified</th>
<th>Type</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>EasyNetMonitor</td>
<td>11/5/2013 5:25 AM</td>
<td>File folder</td>
<td></td>
</tr>
<tr>
<td>recover</td>
<td>11/21/2012 1:34 PM</td>
<td>File folder</td>
<td></td>
</tr>
<tr>
<td>lsak11072014.zip</td>
<td>7/16/2014 8:34 AM</td>
<td>Compressed (zipped) Folder</td>
<td>5,561 KB</td>
</tr>
<tr>
<td>Default.rdp</td>
<td>12/20/2012 8:09 AM</td>
<td>Remote Desktop Connection</td>
<td>0 KB</td>
</tr>
<tr>
<td>pgadmin.log</td>
<td>4/11/2013 10:15 AM</td>
<td>Text Document</td>
<td>2 KB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Date modified</th>
<th>Type</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bag-info.txt</td>
<td>1 KB</td>
<td>No</td>
<td>1 KB</td>
</tr>
<tr>
<td>bagit.txt</td>
<td>1 KB</td>
<td>No</td>
<td>1 KB</td>
</tr>
<tr>
<td>manifest-md5.txt</td>
<td>1 KB</td>
<td>No</td>
<td>1 KB</td>
</tr>
<tr>
<td>tagmanifest-md5.txt</td>
<td>1 KB</td>
<td>No</td>
<td>1 KB</td>
</tr>
</tbody>
</table>
Thank You