Do You have Digitized Collections?

Do You Need To Catalog And Route Those Collections!

Then Why Not Consider...

LadyBird!
Background and Current Status

- The digital collections at Yale library are a wide-ranging aggregation of digitized materials from the Library’s traditional analog holdings augmented by digital media purchased from vendors.

- Setting aside the library’s commercially licensed online books, serials and music, the content is predominantly image-based but the digital collections also include texts, datasets, audio, video and other formats suitable for the source material.

- Principal contributors of digitized images include: Beinecke, Classics, Divinity, Drama, Lewis Walpole Library Manuscript and Archives, Maps, Medical Historical Library, and Visual Resources. Principal contributors of digitized texts include: AMEEL, the Sumner-Smith Collection, the Microsoft project ebooks, and the Yale Daily News.

- These digital collections are created in and reside in several systems and are produced and supported by separate tools, workflows and staff.
The library’s digital collections rest on an ad hoc arrangement of silo software programs that cannot support the weight of expanding collections and inhibit meaningful development of enterprise tools, workflows, and staffing.

In this disconnected landscape, it is impossible to rationalize technological infrastructure, implement standardized metadata schema and asset production, establish uniform quality control, provide a stable cataloging environment and ensure consistency and integrity across the full array of digital assets.

Digital collection managers are hampered by a lack of transparency in the technological organization of these collections as well as a clear path to expansion, and this situation limits the exchange of ideas among staff in a rapidly changing environment.

The absence of an integrated management system with an active user base hinders departments on the threshold of initiating a digital collection. As a result, patrons struggle to discover digital resources.
Solution

- In the spring of 2010, Michael Friscia, James Terray and Rebekah Irwin initiated development of a single comprehensive digital workflow tool with metadata editing features that supports the Library’s multiple digital collections, workflows, and user interfaces called LadyBird.

- LadyBird provides a vital piece of the core infrastructure for collection development and access.

Implementation

- Although only at version 0.3, LadyBird is currently in production use in Beinecke, Manuscripts and Archives, and the Visual Resource Collection. Since May 2011, the VRC has processed more than 12,000 images and associated metadata in LadyBird, and Beinecke processed 3233 new metadata records.

- We are developing, testing, and using the range of LadyBird cataloging features and options, assessing staff feedback, and demonstrating the advantages of this new program as the key tool in a complex array of digital collections, tools and systems.
Principal LadyBird features:

- A robust cataloging tool suitable for complete object description and associated administrative metadata for both simple and complex objects.
- Advanced bulk metadata import and export processing.
- Harvesting metadata from ORBIS and archival finding aids.
- Authority control.
- Bulk digital media processing for a range of file formats.
- Automated transfer of media files and associated metadata to archival storage.
- Flexible workflow options.
- Administrator and staff customization options.
- Shared user documentation.
- And the power to deliver content to multiple user interfaces, such as the Yale library’s Digital Collections interface and ODAI’s cross-collection search.
- The system promotes interaction among collection managers, encourages new participants, and puts in place the foundation on which expanding collections and future technology infrastructure will be built.
An extensible architecture to publish objects produced from many sources for staff use and public services.
Ladybird Technical Fact Sheet:

- Programmed in Microsoft .net/c#
- Version 2.0 and 4.0
- Microsoft SQL Server 2008 R2 for database
- 517,098 digital objects currently in the system
- 617,490 additional digital objects to be imported by July 2012
Ladybird Applications:

- Web application, considered the main face to Ladybird, includes publicly accessible pages to be used for product description.

- SharePoint site for user/technical documentation.

- Job Processing engine - the heart of ladybird importing.

- Export Processing engine - the heart of ladybird exports.

- Bag creation engine - background process that generates bags ready for transfer.

- Heartbeat monitor - background process that checks the health of backend systems.

- Application Cleanup systems - used for general automated system maintenance.

- Development environment.

- Complete with separate database from production environment.
Ladybird User Focused:

- Record deletion and “un-deletion” in GUI interface.
- Authority Controlled Vocabulary and free text metadata manipulation tool, aka Ladybird Toolbox (beta).
- Complex object support.
- Complete integration with EAD metadata import to Ladybird (or preferably XC connection).
- User documentation and training materials.
Ladybird Milestone Requirements:

- Version control system in place for source code.

- Project management software implementation for formal task tracking/deadline scheduling/prioritization (Basecamp).

- Formal bug tracking software to allow progress/reporting/transparency to user community (Jira).

- Open Community focus for product direction from membership outside of Yale.

- Publishing complete (web presentation, preservation and alternative project specific systems).

- Technical documentation for code base.

- Ladybird Management console/tools, aka Birdfeeder.

- Integration of a Library Cloud for digital asset staging/short term storage, aka Bird Nest.
Ladybird Future Development:

Future Development:

- Support for transcripts and WYSIWYG interface for OCR/Transcript text.
- Publishing support for Digital Collections.
- Publishing support for the DAM.
- Bulk and individual object deletion/retrieval via Web interface.
- Additional support for other complex object types.
- Simplified and highly automated import/ingest processes.
- OCR integration.
- Open Source conversion to Java/PHP/MySQL.
LADYBIRD A VISUAL TOUR
### Job List

**ID** | **Job Folder** | **Count** | **Date Start** | **Date Complete** | **User Name** | **Status**
--- | --- | --- | --- | --- | --- | ---
2205 | View Set | 5 | 10/24/2011 4:40:00 PM | 10/20/2011 4:40:01 PM | Sarah Kraus |  
2263 | View Set | 5 | 10/24/2011 4:24:00 PM | 10/20/2011 4:24:01 PM | Sarah Kraus |  

Enter new Job Name: [Merged Selected Jobs] [Delete Selected Jobs]
Dashboard Highlight: Search
Dashboard Highlight: Export Options

Select the fields that will appear when you export. If none are selected ALL will be exported. [Clear checkboxes]

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filename</td>
<td>127</td>
</tr>
<tr>
<td>Accession Auto number</td>
<td>177</td>
</tr>
<tr>
<td>Accession Material</td>
<td>176</td>
</tr>
<tr>
<td>Accession number</td>
<td>59</td>
</tr>
<tr>
<td>Barcode</td>
<td>105</td>
</tr>
<tr>
<td>Box</td>
<td>60</td>
</tr>
<tr>
<td>Call number</td>
<td>58</td>
</tr>
<tr>
<td>Caption</td>
<td>74</td>
</tr>
<tr>
<td>Citation</td>
<td>156</td>
</tr>
<tr>
<td>Classroom use</td>
<td>108</td>
</tr>
<tr>
<td>Coordinates</td>
<td>97</td>
</tr>
<tr>
<td>Creator</td>
<td>69</td>
</tr>
<tr>
<td>Culture</td>
<td>94</td>
</tr>
<tr>
<td>Date returned to Bernacle</td>
<td>157</td>
</tr>
<tr>
<td>Date to Cataloging</td>
<td>156</td>
</tr>
<tr>
<td>Date created</td>
<td>79</td>
</tr>
<tr>
<td>Date, depicted</td>
<td>80</td>
</tr>
<tr>
<td>Date, key</td>
<td>61</td>
</tr>
<tr>
<td>Digital format</td>
<td>157</td>
</tr>
<tr>
<td>Donor</td>
<td>158</td>
</tr>
<tr>
<td>Edition</td>
<td>75</td>
</tr>
<tr>
<td>Ephemera</td>
<td>172</td>
</tr>
<tr>
<td>Folder</td>
<td>61</td>
</tr>
</tbody>
</table>
New Object
Projects Quick Sand Box (cid-31: pld-3)

**Metadata**

<table>
<thead>
<tr>
<th>Catalogue 1</th>
<th>Record modified date</th>
<th>Record ID</th>
<th>Local record ID 1</th>
<th>Local record ID, other 1</th>
<th>Call number 1</th>
<th>Accession number 1</th>
<th>Box 1</th>
<th>Folder 1</th>
<th>Host, Creator 1</th>
<th>Host, Title 1</th>
<th>Host, place of creation 1</th>
<th>Host, publisher 1</th>
<th>Host, date 1</th>
<th>Host, edition 1</th>
<th>Host, note 1</th>
<th>Creator 1</th>
<th>Title 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Choose the following option: 

Submit Form
New Object Highlight: Structure
### New Object Highlight: Publish

<table>
<thead>
<tr>
<th>Setting</th>
<th>Option 1</th>
<th>Option 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use Custom Settings</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Publish to DAM</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Publish to DL</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>DL UNC Size 0 (96px)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL UNC Size 1 (192px)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL UNC Size 2 (384px)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL UNC Size 3 (768px)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL UNC Size 4 (1536px)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL UNC Jpeg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Publish Parent Data</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>UNC for Original Media</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DL Unique ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAM Repository</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>DAM Object Model</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>DAM Security Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAM Bag Prefix</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Submit Publish Settings for THIS OBJECT**
### Project Settings

**Label**
Quick Sand Box

**"Not found" Image**
LadyBird default

**Uncataloged PID or exempt list**
127.152.39.103.157

**Publish to DAM**
- No
- Yes

**Publish to DL**
- No
- Yes

**DL UNC Size**
- Size 0 (90px)
- Size 1 (192px)
- Size 2 (384px)
- Size 3 (768px)
- Size 4 (1536px)

**Publish Parent Data**
- No
- Yes

**UNC for Original Media**
wilcoy/lovespnp.png

**Digital Collections Collection ID**
(\{\{d-1025.1\}\} Yale Center for British Art)

**DAM Repository Code**
N/A

**DAM Object Model**
bibliographicObject

**DAM Security Model**
YUL-Work-in-Progress-Security-Policy

**DAM Bag Prefix**

**Error Email Contact**
michael.fiscio@yale.edu

**Submit Project Settings**
### Project Settings Highlight: Field Selection

Select the fields to display for this project.

<table>
<thead>
<tr>
<th>Handle</th>
<th>Value</th>
<th>Group Label</th>
<th>FID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cataloger</td>
<td>Metadata</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Record modified data</td>
<td>Metadata</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>Record ID</td>
<td>Metadata</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Local record ID</td>
<td>Metadata</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Local record ID, other</td>
<td>Metadata</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Call number</td>
<td>Metadata</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>Accession number</td>
<td>Metadata</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Box</td>
<td>Metadata</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Folder</td>
<td>Metadata</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Host, Creator</td>
<td>Metadata</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>Host, Title</td>
<td>Metadata</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Host, place of creation</td>
<td>Metadata</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>Host, publisher</td>
<td>Metadata</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Host, date</td>
<td>Metadata</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Host, edition</td>
<td>Metadata</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Host, note</td>
<td>Metadata</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Creator</td>
<td>Metadata</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Title</td>
<td>Metadata</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>Title, alternative</td>
<td>Metadata</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>Metadata</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Caption</td>
<td>Metadata</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Parts scanned</td>
<td>Metadata</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Edition</td>
<td>Metadata</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>Place of origin</td>
<td>Metadata</td>
<td>77</td>
</tr>
</tbody>
</table>
### Project Settings Highlight:

#### Project Status

<table>
<thead>
<tr>
<th>Order</th>
<th>Label</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>edit 1</td>
<td>New</td>
<td>delete</td>
</tr>
<tr>
<td>edit 2</td>
<td>Prepare/Add Images</td>
<td><a href="mailto:michael.fricas@yale.edu">michael.fricas@yale.edu</a></td>
</tr>
<tr>
<td>edit 3</td>
<td>Prepare/Add Metadata</td>
<td><a href="mailto:michael.fricas@yale.edu">michael.fricas@yale.edu</a></td>
</tr>
<tr>
<td>edit 4</td>
<td>Verify and Publish this Job</td>
<td><a href="mailto:robert.carlucci@yale.edu">robert.carlucci@yale.edu</a></td>
</tr>
<tr>
<td>edit 5</td>
<td>Complete</td>
<td>delete</td>
</tr>
</tbody>
</table>
Tracking And Log Out
LADYBIRD A CATALOGING TOOL
PURPOSE

- Preservation.
- Cataloging.
- Streamline Workflows.
- Present and Share collections on the web.
REPLACES

- Portfolio
- Luna Studio (for image processing)
- Other image processing software
- Homegrown Web Apps
- “One-off” catalog tools
SECURITY

- CAS authentication.
- Local authentication (non-CAS).
- Field level permission on collection, project, object and metadata.
- Group and role based security model.
FUNCTIONALITY

- Single metadata schema
  - Exports to the DAM Cultural and Bibliographic models
- Rich security models
  - Permits limited read/write for students
- Version control
  - Virtually anything can be rolled back to a prior version
- Easy to use
- Import/export/re-import
**Hierarchy**

- **Collection** (VRC, MSSA, Divinity)
  - **Project** (Silk Road, Glee Club, Eikon)
  - **Object** (digitized book, manuscript, image)
    - **Asset** (metadata record, transcript)
    - **Asset** (image, audio, video)
    - **Asset** (structure)

**Catalog Tool**
HOW DOES ALL THIS WORK?
Drop Folder

- Windows File Share
- Used for import or export
- Works with Web Interface
- Fastest method for cataloging
- Can do almost everything the web interface can do
Web Interface

- Required for some functions
- Customizable
- Screens for
  - Locating objects
  - Editing metadata
  - Editing complex structures
  - Viewing technical metadata
- Publishing to:
  - DAM
  - Digital Collections
  - Other locations set by project
Client Application

- Windows Installer
- Automatic updates
- Small audience
- Handles complex operations not available in web interface or drop folder
- Works with Web Interface
Processing

- Enforces metadata field restrictions
- Generates derivative image files for web presentation
- Stores files on route to ODAI
Publishing

- Communicates with DL interfaces
- Communicates with SQL database
- Sends objects to ODAI
  - DAM
  - Fedora
- Handle generation
WORK FLOWS SIMPLIFIED
Workflow is a sequence of connected steps that encompasses every job or operation of a person or group of persons, an organization, or one or more simple or complex mechanisms. Workflow is the outgrowth of policy, efficiency and availability of resources.

Workflow is central to all tasks. However simple, all jobs, operations and organizations have a workflow.
What is your workflow?

Understanding your departmental needs

Understanding your interdepartmental workflow is important - as it will allow you to make better use of tools such as LadyBird or perhaps even restructure your current workflow to allow for greater productivity.

A LadyBird Workflow encompasses image processing, data processing, cataloging, data review, image review, and use.
What is a LadyBird Workflow?

When thinking about your interdepartmental workflow consider the following:

- Does your content have a review process?
- Are there many individuals working on a single project or is each project handled from start to finish by one individual?
- Are your images made available before the creation of metadata or after?
- Do you generally process content in batches or do you catalog items one at a time?

Regardless of your interdepartmental workflow LadyBird will simplify the process allowing for greater output at an increased speed!
A SAMPLE LADYBIRD WORKFLOW
Step 1. Create a Job Folder

- Ladybird maintains an import folder for each project.
- The name of the new folder will be attached to all the TIF images so that they can be found together later on as a group.
- It is safe to include as many files for import you may need right at the top level.
Step 2. Create a Spread Sheet

- Spread Sheets are only necessary for large sets of metadata or images.
- The naming convention used – serves as a map to your Job folder.
Step 3. Import

- Under the New Objects Tab select Import, then browse and select your spreadsheet.
- Select Import From Source.
Step 4. Process the Job Folder

- Processes are executed on a first in first out basis so that each job is processed in the order in which they are received. Depending on the size of the job, it may take 30 minutes to several hours to complete.
- You will get an email notification when the job is completed. Once completed it will also appear in your Dashboard tab under Job List.
- To continue on to step 5 Click "View Set"
Step 5. Cataloging

- Once you select "View Set" LadyBird brings you to the Search tab where the newly imported thumbnails are shown.

- At this point you have two options for entering Metadata.
  - Option No. One: You can click into a single image and enter metadata
  - Option No. Two: Select the Export As Job link, receive an excel document, add your metadata to that excel document and then re-import that excel document into LadyBird.
Step 5. Option No. One

- Cataloging in the Web Interface
  - Click a thumbnail of any image and you will then be redirected to the New Object / Overview tab. Select the Metadata tab - Within this tab you can edit the metadata for the image.
  - After clicking Submit you will then be sent to the next image in the job. At the top of the page you will see navigation to move between images in the set.
Step 5. Option No. Two

- Cataloging Using Spreadsheets
  - In this step rather than clicking the thumbnail (as in step 5 option No one) you would click the Export As Job link. Once your export is processed by LadyBird you can then fill in the metadata columns using Excel.
  - When done, save the file and Import via the New Objects Tab (see step three).
Step 6. Quality Control

- After you complete the cataloging you should double check your work to make sure that the metadata record matches the image file. This is more important if you used the spreadsheet method. To perform QC go to the Job List Tab and choose the option to View Set (see step 4).

- You can click into individual records or export as a job to an excell document and replace TIF images or metadata content. (**Note to LadyBird Group: it would be nice if you could also choose a report option that produced a report with picture and meta info**).
Step 8. Publish

- Now that your job has been processed you are ready to send the materials to the DAM and to the web. Each project has predefined publishing rules in place to make this step simple. You would then go to the job list and use the "publish" link offered. This will lead you to a page that will indicate any possible errors with your publish. Errors may include unauthorized Authority Control values, multiple values in fields that do not allow repeats or missing values in required fields (Title and Key Date). If everything in the job is ready to be published, click the "publish now" button. As with everything else, your publish will be put into a publishing queue. Publishes are picked up every fifteen minutes and depending on the predefined publishing settings, your new records should appear on the web in about a half hour. Materials sent to the DAM go into a different queue. We are unable to provide an estimate for the length of time it will take for an image to get into the DAM. Be prepared for this process to take several days and possibly weeks. Note that you will receive an email after the materials have been put on the web but not when they have gone into the DAM.
Simplified Workflow By LadyBird

**Beginning with OIDS**

- Import... a large collection via a spreadsheet or process the collection one at a time manually.

**Beginning with Metadata**

- Catalogue... Single images within LadyBird or large sets via an excel spreadsheet.

**Beginning with Images**

- Publish... the collection or continue to work within on the collection in LadyBird... Or export your work and continue to work via a spreadsheet. It is up to you.
LadyBird!

Is Easy To Use!
Fast, Accurate And Fun!

LadyBird Might Just Be, The Answer To All Your Cataloging Needs!

TRY IT TODAY!

http://office.microsoft.com
For Questions Contact:

Michael Friscia:
michael.friscia@yale.edu

Matthew Beacom:
matthew.beacom@yale.edu

Caroline Caizzi:
caroline.caizzi@yale.edu