Collection Management and Creation Strategies for UC Special Collections and Archives:
Summary and Features Matrix

Prepared by CDL Digital Special Collections
October 2009 (rev.)

Context

This document seeks to identify content management and creation systems that UC special collections and archives departments can leverage to locally manage and create digital assets.[1] It is primarily aimed towards UC repositories that either do not have a strategy in place -- or are seeking to replace what they currently use with a more robust solution. The systems that are profiled include both commercial and open-source applications. They range from tools that have been or are currently used by various UC- and non-UC repositories (including solutions that the CDL has endorsed or developed) -- to tools that aren't currently implemented, but may be worth keeping an eye on and assessing further.

Instead of explicitly recommending particular solutions, it seeks to summarize key functional aspects of particular systems (see **Appendix A** and **B**), including licensing type, degree to which the product can be customized by developers, content management and/or creation features, etc. In particular, it also highlights whether the application can generate EAD- and METS-based outputs, which may be submitted to CDL for preservation and access services. Below are some general questions that may help guide your use of the summary table:

- Is it important for you to be able to publish your content to the web?
  - If so, see **Appendix A**: Archon, CollectiveAccess, CONTENTdm, Cuadra STAR/Archives, DigiTool, DSpace, Eloquent Archives, Fedora, ICA-AtoM, Past Perfect, ResCarta Toolkit

- Are you only interested in open source solutions?
  - For collection management systems -- see **Appendix A**: Archivists' Toolkit, Archon, Collective Access, DSpace, Fedora, GenDB, ICA-AtoM, ResCarta Toolkit
  - For content creation tools -- see **Appendix B**: OAC EAD Web Templates, XML authoring tools

- Are you seeking a comprehensive collection management solution, or tools to be able to create EAD collection descriptions or METS digital objects only?
  - For collection management systems only: see **Appendix A**
  - For content creation tools only: See **Appendix B**

- Is the option to create finding aids (that can be ultimately exported as EAD) important to you?
  - If so, see **Appendix A**: Archivists' Toolkit, Archon, Cuadra STAR/Archives, Eloquent Archives, GenDB, ICA-AtoM, RE:Discovery Proficio
  - See also **Appendix B**: OAC EAD Web Templates, XML authoring tools

---

[1] The following terminology is used throughout this document:

- **"Collection Management Systems":** integrated system for managing collection holdings. Typical features may include:
  - Provides collection management features (e.g., tracking locations, usage statistics, processing status, etc.); supports accessioning and deaccessioning, work plans, description, etc.
  - Reuse of data for multiple outputs and purposes (e.g., descriptive info. for an item can be used for accession reports or to create a digital object for display)
  - Supports importing of collection or digital object data (e.g., legacy data), and outputs to multiple formats (shareable data for aggregators, etc.)
  - Provides web publication functions
  - Supports use of content and data standards

- **"Content Creation Tools":** specifically geared towards the creation of collection guides or digital objects, to be exported in the EAD or METS format.

- Is the option to create digital objects (that can be ultimately exported as METS) important to you?
    - If so, see **Appendix A**: Archivists' Toolkit, CONTENTdm, DigiTool, DSpace, Eloquent Archives, Fedora, GenDB, ResCarta Toolkit
    - See also **Appendix B**: docWorks/METAe, XML authoring tools

For an in-depth features matrix and discussion of solutions that are highlighted here (including suggested criteria for selecting particular applications -- e.g., L. Spiro's report, Chapters 5-6) -- in addition to other applications -- see the following references:

- Council of Nova Scotia Archives. *Archives Management Software Review* (2009): http://www.councilofnsarchives.ca/resources/SoftwareReview.htm
- Spiro, Lisa. *Archival Management Software: A Report for CLIR* (2009): http://clir.org/pubs/reports/spiro/spiro_Jan13.pdf
- Wisser, Katherine M. "EAD Tools Survey" (2005): http://www.archivists.org/saagroups/ead/EADToolsSurvey.pdf

Our summary draws heavily on information compiled by the Council of Nova Scotia Archives and L. Spiro. The CDL Digital Special Collections program does not in any way endorse any product or application discussed in this report; we have attempted to highlight solutions that we feel may be of particular interest to the intended audience of this document. Omission of a particular product does not indicate disapproval. We welcome any comments, suggestions, updates, or corrections to our summary (please send to oacops@cdlib.org).

Summary of "Pros" and "Cons"

Below is a summary of very generalized "pros" and "cons" with different strategies outlined in more detail in **Appendix A** and **Appendix B**.

Collection Management Systems: Open-Source Solutions

Application	Archivists' Toolkit 1.5	Archon 2.21	Collection Space	CollectiveAccess 0.55-1	DSpace	Fedora	GenDB	ICA-AtoM	ResCarta Toolkit
PROS	* Archival collection tracking/ management support * Finding aid and digital object creation support * Supports integration of finding aids w/ digital objects * EAD export * METS export	* Archival collection tracking/ management support * Finding aid and digital object creation support * Supports integration of finding aids w/ digital objects * Integrated web publication features * EAD export	* Still in development	* Archival collection tracking/ management support * Finding aid creation support (partial) and digital object creation support * Integrated web publication features	* Digital object creation support * Integrated web publication features * METS export	* Digital object creation support * Integrated web publication features * METS export	* Supports integration of finding aids w/ digital objects * EAD export * METS export	* Finding aid and digital object creation support * Supports integration of finding aids w/ digital objects * Integrated web publication features * EAD export	* Digital object creation support * Integrated web publication features * METS export
CONS	* Limited digital object management/ manipulation capacity at this time (e.g., viewing, batch metadata entry, etc.) * No integrated web publication features	* No METS export	* Still in development	* No integration of finding aids w/ digital objects (potential for upcoming release?) * No METS and EAD exports (although planned for upcoming release)	* No archival collection tracking/ management support? * No finding aid/EAD creation capacity (displays EAD only)	* No archival collection tracking/ management support? * No finding aid/EAD creation capacity (displays EAD only)	* No archival collection tracking/ management support? * No integrated web publication features	* No archival collection tracking/ management support? * No METS export (although planned for upcoming release)	* No archival collection tracking/ management support? * No finding aid/EAD creation capacity (displays EAD only) * No EAD export

Collection Management Systems: Commercial Solutions

Application	CONTENTdm 5.0	Cuadra STAR/Archives	DigiTool 3.0	Eloquent Archives	Past Perfect 4	RE:Discovery Proficio 8
PROS	* Digital object creation support * Integrated web publication features * METS export (when used in conjunction with separate 7Train application)	* Archival collection tracking/ management support * Finding aid and digital object creation support * Supports integration of finding aids w/ digital objects * Integrated web publication features * EAD export	* Digital object creation support * Integrated web publication features * METS export	* Some collection tracking/ management support * Finding aid and digital object creation support * Supports integration of finding aids w/ digital objects * Integrated web publication features * EAD export	* Some archival collection tracking/ management support * Finding aid and digital object creation support * Integrated web publication features	* Some archival collection tracking/ management support * Finding aid and digital object creation support * Integrated web publication features (requires RE:Discovery for the Internet module)
CONS	* METS export requires post-processing (using 7Train) * No archival collection tracking/ management support * No finding aid/EAD creation capacity (displays EAD only)	* No METS export	* No archival collection tracking/ management support? * No finding aid/EAD creation capacity (displays EAD only)	* Partial METS export capacity, but requires post-processing	* Integration of finding aids w/ digital objects -- capacity is not known * No EAD export * No METS export	* Integration of finding aids w/ digital objects -- capacity is not known * No METS export

Content Creation Tools

Tool	docWorks/METAe	OAC EAD Web Templates	XML authoring tools (commercial products)	XML authoring tools (open source)
PROS	* Generates METS * Possible validation support?	* Easy to use; simple web form (doesn't require detailed knowledge of EAD)	* Some of these tools can be used in conjunction with the popular EAD Cookbook, which provides encoding and display templates. * Can establish templates to facilitate data entry and compliance with best practice guidelines * Easy to use (similar to word processor) * Generates EAD and METS; also generates other outputs (e.g., HTML for the web or for conversion to PDF) * Can be configured to validate on-the-fly	* Some of these tools can be used in conjunction with the popular EAD Cookbook, which provides encoding and display templates. * Can establish templates to facilitate data entry and compliance with best practice guidelines * Easy to use (similar to word processor) * Generates EAD and METS; also generates other outputs (e.g., HTML for the web or for conversion to PDF) * Can be configured to validate on-the-fly
CONS	* Manual EAD or METS creation	* Manual EAD creation * Any editing of the EAD file requires use of a different encoding tool (and some experience with EAD) * Cannot be used to generate detailed container lists * Cannot be used to save files in process; "one off" process * Requires separate EAD validator (OAC BPG EAD Desktop Validator)	* Manual EAD or METS creation	* Manual EAD or METS creation

What are Calisphere/OAC Contributors Doing Now?: Some Current Strategies

Case Study #1A: UC Berkeley's GenDB service

This scenario is drawn from the Library of Congress-funded "California Cultures" digitization project (2000-2005), in which multiple UC campus special collections and archives participated. In this scenario, campuses utilized a single, shared open source collection management system hosted by UC Berkeley's Digital Publishing Group (DPG) -- GenDB -- in order to create and maintain digital objects. Various different content creation tools were utilized by some campuses to specifically generate EAD finding aids. To integrate the digital objects with the EAD finding aids, the system users post-processed their EAD files to include references to the digital objects. Some UC special collections and archives units currently use this strategy.

Calisphere/OAC contributors have a few potential options for using GenDB:

1) **Host locally:** Under this arrangement, institutions would install and manage GenDB locally.
2) **Utilize a hosting service -- UC Berkeley DPG:** UC Berkeley's DPG may be able to provide GenDB services for interested UC campus repositories; for more information, contact DPG (http://www.lib.berkeley.edu/cgi-bin/dpg_contact.cgi).

The OAC EAD Web Templates are freely-available for use; for more information see http://www.cdlib.org/inside/projects/oac/toolkit/templates/.

Summary

Digital objects solution	Finding aids solution
• GenDB	• OAC EAD web templates • XML authoring tools (commercial products or open source)

Case Study #1B: UC Berkeley's GenDB service

This scenario illustrates a current variation of the strategy that was employed in the California Cultures project (Case Study #1A). In this case, UC Berkeley's GenDB is now used as a single, shared open source collection management system that also has the capacity to generate EAD finding aids.

See Case Study #1A for GenDB usage options.

Summary

Digital objects solution	Finding aids solution
• GenDB	• GenDB

Case Study #2: Archivists' Toolkit

This is similar to Case Study #1B in that a single collection management (that also has the capacity to generate EAD finding aids) is being utilized. The Archivists' Toolkit (AT) is a freely-available open source application, and is implemented using a client-server arrangement: it requires a MySQL database to manage records, and one or more client applications to interact with the database. Calisphere/OAC contributors have a few potential options for using the AT:

1) **Host locally:** Under this arrangement, institutions would install and manage both locally.
2) **Utilize a hosting service -- CDL Digital Special Collections:** The CDL Digital Special Collections program is also currently investigating hosting of the AT for use by our OAC contributors: under this arrangement, the CDL would host and maintain a single, shareable instance of an AT MySQL database -- contributors would only need to download the AT client, in order to use the service. By providing this service, our goal would be to minimize costs associated with creating or hosting the AT locally.

For more information about the AT, see http://www.archiviststoolkit.org.

Summary

Digital objects solution	Finding aids solution
• Archivists' Toolkit	• Archivists' Toolkit

Case Study #3: UC institution-specific customized solutions

In this scenario, the institution created a custom collection management solution to fit their local needs, using commercial or open source products such as FileMaker, Microsoft Access, MySQL, etc. The same application -- or a separate application is used to generate EAD finding aids.

Summary

Digital objects solution	Finding aids solution
• UCLA DLIS • UCSD DAMS	• UCLA DLIS • XML authoring tools (commercial product)

Case Study #4: CONTENTdm

This case is drawn from the Library and Technology Services Act (LSTA)-funded "Local History Digital Resources Project (LHDRP)" (2002-current), in which multiple non-UC (primary public libraries) and UC library special collections and archives have participated. In this scenario, participants utilize a single, shared commercial collection management system hosted by the Califa Library Group -- CONTENTdm -- in order to create and maintain digital objects. The digital objects can be exported and transformed into METS by the CDL, using the open-source 7Train METS Generation Tool (http://seventrain.sourceforge.net/). The OAC EAD Web Templates are used by the participants to specifically generate EAD finding aids. To integrate the digital objects with the EAD finding aids, the system users post-process their EAD files to include references to the digital objects.

CONTENTdm is a commercial application, and is implemented using a client-server arrangement: it requires a central web-based server to manage records, and one or more client applications to interact with the server. Calisphere/OAC contributors have several options for using the AT:

1) **Host locally:** In this scenario, institutions would install and manage both the client and server locally.
2) **Utilize a hosting service -- OCLC (FirstSearch® Base Package):** We recently learned that the UC Libraries may be gaining OCLC's FirstSearch® Base Package in May 2009, as part of the Next Generation Melvyl Pilot supported by OCLC. The base package includes access to and use of CONTENTdm Version 5.0 software. OCLC will host the server, and libraries will be able to implement up to three clients. For more information, see http://0-www.oclc.org.millennium.mohave.edu/us/en/services/brochures/211378usb_whats_new_at_oclc.pdf . Additional details will be forthcoming to the UC Libraries, as the information becomes available.
3) **Utilize a hosting service -- Califa:** Califa also provides use of its CONTENTdm service to its membership on a cost-recovery basis (for server space, staffing/administration, etc.). Califa hosts the server, and libraries implement one or more clients. For more information about membership and fees, see http://www.califa.org/cdm_info.php.

For general information about CONTENTdm, see http://www.contentdm.com/.

Summary

Digital objects solution	Finding aids solution
• CONTENTdm (used in conjunction with 7Train)	• OAC EAD web templates

Case Study #5: DigiTool

Similar to Case Study #4, a commercial collection management system -- DigiTool -- is used as a content management solution for creating and maintaining digital objects. To generate EAD finding aids, separate commercial and/or open source products are used. To integrate the digital objects with the EAD finding aids, the system users post-process their EAD files to include references to the digital objects.

For general information about DigiTool, see http://www.exlibrisgroup.com/category/DigiToolOverview.

Summary

Digital objects solution	Finding aids solution
• DigiTool	• OAC EAD web templates • XML authoring tools (commercial products or open source)