Next-Generation Technical Services (NGTS)

Archivists' Toolkit Recommendations

Final Report

March 2012 Rev. August 2012

Power of Three (POT) #3, Lightning Team #1
David Gartrell, UC Santa Barbara (Co-chair)
Eric Milenkiewicz, UC Riverside
Adrian Turner, CDL (Co-chair)
Brad Westbrook, UC San Diego

1. Executive Summary

Our team has been charged with developing an implementation plan for the coordinated, system-wide usage of the Archivists' Toolkit (AT), in accordance with the recommendations of earlier NGTS efforts. During the winter of 2011, we conducted an assessment of AT usage and needs within the UC libraries, and determined that that many campus repositories are already using the software to some degree. In other words, implementation is already happening. In lieu of a deployment and implementation plan, what we believe is needed is a greater emphasis on the coordination of best practices and the efficient use of the application. The following report offers two key recommendations to facilitate this. For the two cases where campus repositories are not using the AT — but where implementation may occur in the future and could benefit from assistance with "jump starting" usage — our report additionally discusses considerations involved with implementing the software locally versus using California Digital Library's (CDL) hosted service.

Our two key recommendations are:

Recommendation A: UC AT User Group

- <u>Summary:</u> Establish a special interest all-campus library group of knowledgeable AT users comprised of members across the UC libraries in order to facilitate coordinated system-wide use of the AT and to provide ongoing support for UC units using the AT.
- <u>Timeframe:</u> Summer or Fall 2012; estimate 1 month to complete proposed tasks identified below
- <u>Costs:</u> Resource needs and costs to establish this group and a subsequent online portal are seen as negligible.

Recommendation B: UC AT Training

- <u>Summary:</u> We recommend a "train the trainer" workshop addressed to UC library repositories. A "train the trainer" approach will serve to provide uniform training to UC staff and to cultivate the formation of a cross-campus expert user community. The workshop would be an in-person event, offered on 2-3 consecutive days (basic and advanced features; specific topics to be developed in consultation with the UC AT User Group) in a northern and southern California UC library location. Representative(s) from each UC library campus repository that is currently utilizing the AT, or interested in utilizing the AT, should attend (computer lab size will dictate the number of additional attendees).
- <u>Timeframe:</u> Fall or Winter 2012; estimate 2 months to complete proposed tasks.
- <u>Costs:</u> Relatively minimal costs: in-kind campus and CDL contributions of staff time to support training, and travel costs for attendees to participate in the workshop; in-kind campus contribution of meeting room and local arrangements.

2. Context

Our team has been charged with developing an implementation plan for the coordinated, system-wide usage of the AT, in accordance with the recommendations of earlier NGTS efforts. In order to develop such a plan, our initial step was to survey AT users and potential users within the UC libraries — specifically special collections and archives — to identify how AT is currently being used, what, if any, impediments to use exist amongst non-users, and what sort of assistance or information might be needed by both groups in the adoption and ongoing use of the software.

During the winter of 2011, we collected survey data from representatives of each department of special collections and archives within the UC library system. We asked these key users about current practices, needs, and preferences in archival software use and training. Four primary criteria for the effective implementation of the AT emerged from the survey data: broad staff training in the basics of the AT, with additional, more advanced training for some staff members; resources to facilitate the integration of the AT into local workflows; ongoing support for AT users; and the need for a succinct AT marketing brief.

Our survey revealed that there is already broad usage of the AT by special collections and archives in the UC Libraries; most campuses are already using it to some degree. In other words, the software has already been largely deployed and implementation is already happening — and there is a "critical mass" usage. Below is a list of UC campus libraries currently implementing the AT:

- UCSB Library, Department of Special Collections (locally-hosted)
- UCLA Library, Special Collections (locally-hosted)
- UCD Library, Special Collections (locally-hosted)
- UCR Libraries, Special Collections & Archives (locally-hosted)
- UCSD Libraries, Mandeville Special Collections Library (locally-hosted)
- UCM Library/Digital Assets, Special Collections and University Archives (CDL-hosted)
- UCI Libraries, Special Collections & Archives (locally-hosted)
- UCSF Library, Archives & Special Collections (CDL-hosted)

Below is a list of UC campus libraries that have not implemented the AT:

- UCB Library, Bancroft Library (currently utilizing local systems -- GenDB, MyECMS -- that are already tailored to collection management and processing workflows)
- UCSC Library, Special Collections & Archives (implementation of the AT is planned)

Our recommendations address these two primary use cases:

A. UC campus library repositories that are currently and actively implementing the AT:

In lieu of a deployment and implementation plan, what we believe is needed is a greater emphasis on the coordination of best practices and the efficient use of the application. We propose two recommendations to support this, which we believe can cultivate more effective use of the application throughout the system, with relatively low in-kind costs: 1) bringing together

knowledgeable AT users from across the campuses to facilitate the integration and ongoing support of the AT and, 2) training targeted to UC-specific needs.

B. UC campus library repositories that would like to implement the AT, and could benefit from its use:

For cases where repositories are interested in utilizing the AT and are opting to implement it, we propose that the recommendations outlined above will also promote effective use of the application. Additionally, these cases should take into account CDL-hosting versus local-hosting options:

CDL-hosted AT

This option provides CDL-hosting of the AT if the repository doesn't have the necessary local IT resources or capacity to maintain the AT database. In the CDL-hosting arrangement, the CDL mediates access to the database, and manages software updates, maintenance, data recovery processes, etc. The CDL is aware of two cases (out of a total of 33 hosted instances) where repositories have experienced intermittent disruptions in connections between their client software and the hosted database; it is currently investigating solutions to maintain the connection. Additional details about the hosting service are available on the CDL website.¹

Locally-hosted AT

This option should be considered if the repository has local IT resources and capacity to maintain the database, and wants unmediated control over the software. In a local hosting arrangement, the repository has direct oversight and administration over the database, and is responsible for software updates, maintenance, data recovery processes, etc.

Hosting considerations

In general, there are relatively low costs to hosting the AT. The AT is itself an open-source product and is available to download for free. Hardware and other material costs comprise hosting of one of the three supported back-end databases (MySQL 5.0, Microsoft SQL Server 2005, and Oracle 10g) and maintenance of that host environment. For both the campus-hosted and CDL-hosted deployments, it is estimated that IT staffing costs to support the software are fractional, at the level of less than 1% of 1 IT staff member's time (20-40 hours total) for initial setup and configuration of software. Beyond the initial setup work, ongoing maintenance of the databases is a marginal cost, but we assume that there will be incremental costs based on the addition of more content. IT staff time allocations would rise to accommodate upgrades, migrations, and release packages that particularly pertain to the back-end databases.

Repositories that are currently maintaining locally-hosted instances do have the option to migrate to the CDL-hosted instance; this is a relatively simple process, involving the campus generating a MySQL dump of the locally-hosted database files, and transferring them to CDL. The CDL can then reconstitute the database as a CDL-hosted instance. That said, we do not believe that there are any significant cost savings for institutions to embark on migrating from existing locally-hosted instances to a CDL-hosted instance, given the relatively low ongoing maintenance costs -- and the

4

¹ http://www.cdlib.org/services/dsc/contribute/at-archon.html

fact that upfront costs (initial setup and configuration of the software) have already been encumbered.

Legacy data considerations

Campuses have a differential amount of legacy data -- ranging from print and EAD finding aids, to accession records -- that they may consider migrating into the AT. In cases where campuses have deployed the AT, they have typically conducted a gradual, phased approach for importing legacy data.

The AT supports batch importing of legacy data in particular formats -- namely, accession records in tab-delimited or XML format (where the data is formatted according to the internal AT database schema), finding aids in MARCXML or EAD formats, and digital objects in tab-delimited format or as part of an imported finding aid. These records can easily batch imported into the software, using the AT's built-in import processes -- but repositories may choose to conduct some quality control checking and data cleanup of the imported records, especially if the source files lack particular standardized descriptive data. Assuming that the files are correctly formatted and/or valid MARCXML or EAD, the staff time required to run the import process should be negligible. The primary consideration when importing large amounts of files is the AT software processing time: large files will take more time to import. We recommend that batch processes are run overnight.

Migration of other kinds of legacy collection management information (including accession records, finding aids, donor files, box lists and inventories, in other formats, such as MS Word documents, PDF files, etc.) into the system may require much more local staff intervention, depending on the state and form of the source data -- and the degree to which individual repositories would like to consolidate this information in the AT. In short, there is no "one size fits all" solution, and other kinds of data will require analysis and some degree of local resource allocation to the process. Given that most collection management information will be highly local and specific to the repository, we do not think this type of work can be easily outsourced -- and will require an investment of local resources. The proposed UC AT User Group can serve as a locus for consultation and collective problem-solving of specific issues pertaining to importing of local data.

3. Recommendation A: UC AT User Group

Summary of Key Findings

- A common theme emerged from our survey pointing to the need for increased communication surrounding the AT and ongoing support for AT users at both the campus and system-wide levels.
- Respondents also expressed a need for greater awareness of the software itself including the AT's overall functionality and benefits, as well as for assistance with integrating the AT into local workflows and with other systems.

Recommendation

<u>Proposal</u>

We propose establishing a UC AT User Group comprised of members across the UC libraries in order to facilitate coordinated system-wide use of the AT and to provide ongoing support for UC units using the AT. The group would also be responsible for adapting practices to the forthcoming ArchivesSpace software (and would be well-suited to conduct early testing and evaluation of the product).

The UC AT User Group will be formally charged and established as a special interest all-campus library group. We recommend that the group has a reporting line to the Heads of Special Collections (HOSC); we additionally recommend a member of the Cataloging and Metadata Common Interest Group (CAMCIG) is appointed to serve on the group, as a formal liaison. It is expected that the group will require few, if any face-to-face meetings, as most of its work can be accomplished using email, wikis, and other communication tools.

Membership on this group should consist of archivists, librarians, curators, and other stakeholders in the UC system (primarily but not exclusively UC Libraries' staff) that use the AT to satisfy their responsibility for the description of archival materials. Each campus library should designate a member from its campus to serve on this group. That person, in addition to system-wide projects, would be responsible for coordinating AT related efforts on their campus. Such campus related efforts would not necessarily be restricted to the libraries, and could extend to other campus units that are using or are interested in using the AT for the description of archival materials. Extending the concerns of the expert group beyond the campus libraries will help to establish a more uniform set of practices for using the AT and for describing archival materials throughout the UC system. Uniformity of this kind will help to decrease the cost of archival description and improve the utility of the Online Archive of California (OAC) database of archival finding aids.

The specific responsibilities of this group will include:

- Promote AT use at the campus and system-wide levels by:
 - o Identifying stakeholders across the system
 - Sharing information to conserve resource expenditures
 - o Sharing third-party or UC-developed resources and tools supporting use of the AT
 - o Promoting best practice guidelines for the AT system-wide
 - Optimizing workflow(s) at the campus/unit levels (including the integration of AT record exports in OAC, Melvyl, and WorldCat); sharing solutions and techniques for integrating the software into local workflows
 - o Developing and providing AT training materials
 - Tracking and evaluating AT-related projects, especially the ArchivesSpace project. The UC AT User Group would be particularly well-situated to conduct early testing of the forthcoming ArchivesSpace software.
 - o Identifying collaborative opportunities throughout and beyond the UC system
 - o Serving as an information conduit between AT project staff and UC AT users
 - Serve as a resource for information about the AT, for UC campus library staff (e.g., supply informational or marketing briefs about the software, to facilitate an understanding of the software's functions)
 - Conducting ongoing assessment of AT usage system-wide; we suggest that the group could reprise the survey work conducted by NGTS POT3 LT1 at a later time, in order to comparatively evaluate usage information

It is anticipated the UC AT User Group would satisfy some of its responsibility by creating and managing an online portal that would provide ongoing support for UC AT users. This portal would include such things as best practice guidelines, information on the integration of AT record exports with existing access/discovery systems, simple tutorials, user stories, and FAQs. Appropriate links to the broader AT community and forums would also be established to support the overall use and adoption of the AT.

A draft charge is provided in Appendix A of this document.

Target Start Date

• Summer or Fall 2012; estimate 1 month to complete proposed key tasks

Key Tasks

- Once recommendation is approved, NGTS Management Team identifies the reporting line and primary UC Library system-wide group that provides oversight for the UC AT User Group; we suggest that this group reports to HOSC, and includes a liaison from CAMCIG.
- Primary UC Library system-wide group finalizes description and charge for the UC AT User Group
- Primary UC Library system-wide group identifies potential members to serve on the UC AT User Group, invites participants, and finalizes membership
- UC AT User Group receives formal charge and is instantiated

Resource Needs and Cost Analysis

Resource needs and costs to establish this group and a subsequent online portal are seen as negligible. The majority of costs involved would be for staff to participate, as with existing UC Libraries' special interest all-campus library groups (such as the Acquisitions Common Interest Group, Digital Reference Common Interest Group, etc.). Additional costs would include the technical framework for the online portal and its management/maintenance as well as communications costs for group meetings.

Overall costs of utilizing the AT are expected to be lower via a system-wide approach to sharing expertise, than they would if each campus were to utilize AT instances independently of other campus or system units. This reasoning dovetails with the reasoning for AT expert groups in other locations such as the Northwest Digital Archives and Harvard University.²

In summary, we anticipate relatively small resource costs to implement this proposal, which we think will ultimately help us avoid larger costs (that would be incurred if the UC Libraries' continue to lack shared and broad expertise of this application).

Benefits

- Promote standardized and efficient use of the AT throughout UC (campus and system-wide levels)
- Optimize AT workflows throughout UC (campus and system-wide levels)
- Provide better and more consistent data for dissemination via OAC

² For information about the NWDA Archivists' Toolkit Working Group, see http://orbiscascade.org/index/nwda-archivist-s-toolkit-working-group. For information about the Harvard University Archivists' Toolkit Working Group, see: https://mailman.ucsd.edu/pipermail/atug-1/2007/000178.html

- Reduce costs of archival description
- Provide UC AT users with a support network that facilitates communication and sustained support of UC AT users
- Broad, cross-campus expertise with the software will facilitate implementation of software upgrades -- as well as transitioning to the forthcoming ArchivesSpace application
- Group can serve as a single voice for comments/feedback to the software developers, to facilitate updates and enhancements to the application to meet UC campus needs.
- Builds on the existing UC Libraries' special interest all-campus library group and working group structures, which have been successful models for cross-campus collaboration

Dependencies (Beyond Approval of Recommendation/Resource Allocation)

• Investment of staff resources by UC campus libraries, to serve on the group

Barriers/Obstacles to Implementation

- Funding and support of an online portal
- Staffing/resource costs associated with the formation of a special interest all-campus library group
- Locating potential members for the group at campuses currently not using the AT

4. Recommendation B: UC AT Training

Summary of Key Findings

- We received responses from all UC special collections and archives represented through HOSC; 9 out of 10 surveyed departments indicated the need for training in some aspect of AT for a total of forty-five staff members.
- The majority view was that in-person training on each campus (80% of respondents) would be preferred to a webinar format or to sending multiple staff to a regional training workshop.
- The most indicated needs for training were in "Basics of Archivists' Toolkit," "Imports and data cleanup," and "Plug-ins," with 8 out of 10 departments expressing interest; additionally, 7 out of 10 departments expressed interest in "Digital objects" and "Reports"

Existing Options

The following table summarizes AT training options that are currently available, or will be offered in the near term by third-party organizations and the AT developers:

Option	Description	Considerations	Costs
Society of American Archivists (SAA) continuing education workshop	 Offered on request and with local support (electronic classroom, etc.); open to cosponsorship model for UC-specific audience 2 day in-person workshop SAA covers registration, logistics with securing a training location, contracting with instructor, etc. Up to 25 attendees per session 	 Offered as part of SAA continuing education program Can be co-sponsored (e.g., if UC wanted to host a workshop for designated attendees) Availability subject to SAA continuing education calendar; not regularly scheduled but only in response to interest at a given locale Attendees must travel to workshop location Content of workshop not customizable 	Approximately \$325 (SAA members)/\$375 (non-SAA members) per attendee for registration, plus travel costs \$5,000 flat fee for co- sponsoring

Society of California Archivists (SCA) continuing education workshop	 Offered on an ad hoc basis; would need to explore co-sponsoring a UC-specific workshop (with extra seats open to SCA membership) 1 day in-person workshop SCA covers registration, logistics with securing a training location, contracting with instructor, etc. Generally up to 25 attendees per session 	 Sporadic availability Availability subject to SCA continuing education calendar; not regularly scheduled but only in response to interest at a given locale Attendees must travel to workshop location Content of workshop not customizable 	Approximately \$180 (SCA member)/\$220 (non-SCA member) per attendee for registration, plus travel costs
Contracting with AT developers	 Offered on a contractual basis Flexible sessions (1, 2, or 3 days) Host covers logistics with securing a training location Generally up to 25 attendees per session, or as the training location allows 	 Three different workshops are available: one-day for AT core functions, two-day for AT basics and some advanced features, and one-day for digital objects Customizable to meet specific local needs No travel for participants 	\$500 per day per instructor (up to \$2000) and travel, lodging, per diem per instructor if necessary. Instructor fee subject to change.
Orbis Cascade Alliance webinars	 Offered on a contractual basis Flexible sessions (1 or 2 days) Host covers logistics with securing a training location Generally up to 25 attendees per session, or as the training location allows 	 Recordings can be consulted on an on-demand and ad hoc basis No travel for participants Content of workshop not customizable 	Anticipated costs are approximately \$100 per attendee

Recommendation

<u>Proposal</u>

We recommend a "train the trainer" workshop addressed to the UC Libraries. Initial training for members of the UC AT User Group from each campus could be provided at two regional locations in southern and northern California, as in-person workshops. A "train the trainer" approach will serve to provide uniform training to UC staff and will also help to cultivate the formation of the UC AT User Group by providing members with an opportunity for in-person discussions.

Rationale

A "train the trainer" approach poses several advantages over other training options currently available:

- It would serve to lower overall costs without diminishing quality
 - o Travel would be restricted to California
 - o Instructor fees would be absorbed by UC and not paid out-of-pocket by each workshop attendee
 - Duration of the workshop could be extended by one day (e.g., a three-day workshop)
 allowing for a quantity of content that otherwise requires attending two different
 workshops if sponsored by SAA or SCA
- It would assure that training is uniform and completed at the same time for all members of the UC AT User Group, and training would not be dependent on the schedules and space limitations of SAA or SCA AT workshops
- It would help to accelerate a UC AT User Group and would foreground concerns and needs that are pertinent to UC campuses
- It would use the same faculty that is used for the SAA or SCA workshops. That faculty has provided over 40 AT one- and two-day workshops across the U.S. since March 2008

For any campus wanting to do so, the "train the trainer" workshop could be augmented by campuses utilizing other existing training options, as previously noted. Campuses may also consider contracting with other experts (e.g., XSLT or JasperReports programmers), if necessary.

Scope

- In-person workshop, offered on 2-3 consecutive days (basic and advanced features; specific topics to be developed in consultation with the UC AT User Group)
- We recommend that the workshop is offered at least once at a northern and once at a southern California UC library location. Computer lab size will dictate the number of attendees.
- Representative(s) from each UC library campus repository that is currently utilizing the AT, or interested in utilizing the AT, should attend. Additional seats would be distributed across the campuses (for other staff to attend), as extra seats allow.
- Content would be highly customized to meet participant training needs
- Workshop objectives
 - o Introduce campus staff to the AT application, covering basic and advanced features
 - o Introduce campus staff to training materials, to adapt for local use

Target Start Date

• Fall or Winter 2012; estimate 2 months to complete proposed key tasks

Key Tasks

- UC AT User Group to coordinate with instructors to identify workshop locations and dates, and confirm local arrangements
- Instructors to develop training materials, in consultation with UC AT User Group
- Host workshops

Resource Needs and Cost Analysis

- Meeting room and local arrangements:
 - o Computer lab. For technical requirements, see http://www.archivists.org/prof-education/logistics.asp#IDACS
 - o Hosting costs absorbed by UC campus library, one in northern California and one in southern California
- Training materials:
 - o To be determined; but likely only printed handouts would be needed
- Participant/attendee time; include time for presentations and information sharing.
- Instructor/convener time:
 - o In-kind support through UCSD, for Brad Westbrook's time as instructor/convener (estimate 40 hours of preparation and 40 hours of instruction including travel)
 - o In-kind support through CDL, for Adrian Turner's time as co-instructor/convener (estimate 40 hours of preparation and 40 hours of instruction including travel)
- Travel costs for instructor:
 - Travel costs provided as in-kind support by UCSD for Brad Westbrook and by CDL for Adrian Turner
- Travel costs for attendees:
 - o Travel costs for participants to be covered by UC campus library, as in-kind support

In summary, we anticipate relatively small resource costs to implement this proposal, which we think will ultimately help us avoid larger costs (that would be incurred if the UC Libraries continue to lack shared and broad expertise with the software).

Benefits

- Promote uniform and timely implementation of the AT system-wide
- In tandem with Recommendation A (UC AT User Group), provides the basis for a sustainable mechanism for each UC campus to provide training at the local level: designees would serve as local expert users and training resources for their local campus processing staff(s).
- Promote standardized and efficient use of the AT throughout UC (campus and system-wide levels)
- Optimize AT workflows throughout UC (campus and system-wide levels)
- Provide better and more consistent data for dissemination via the OAC
- Reduce costs of archival description
- Provide UC AT users with a support network (proposed UC AT User Group) that facilitates communication and sustained support of UC AT users

Survey respondents showed preference for in-person training sessions on each campus. This recommendation offers a hybrid approach, in which a small number of key contacts (UC AT User Group) from each UC campus would be trained to deliver training and support to AT users at their home campuses.

Dependencies (Beyond Approval of Recommendation/Resource Allocation)

• Implementation of Recommendation A (UC AT User Group).

Barriers/Obstacles to Implementation

• Resource constraints pose the biggest obstacle to conducting this workshop

Appendix A: UC AT User Group -- Charge

Charge

The UC AT User Group supports and coordinates system-wide use of the AT. The group is responsible for guiding system practices to the later iterations of the software (and is well-positioned to conduct early testing and evaluation of the products).

Specific responsibilities include:

- Identifying stakeholders across the system
- Sharing information to conserve resource expenditures
- Sharing third-party or UC-developed resources and tools supporting use of the AT
- Promoting best practices for the AT system-wide
- Optimizing workflow(s) at the campus/unit levels (including the integration of AT record exports in OAC, Melvyl, and WorldCat); sharing solutions and techniques for integrating the software into local workflows
- Developing and providing AT training materials
- Tracking and evaluating AT-related projects, especially the ArchivesSpace project. The UC AT
 User Group will be particularly well-situated to conduct early testing of the forthcoming
 ArchivesSpace software.
- Identifying collaborative opportunities throughout and beyond the UC system
- Serving as an information conduit between AT project staff and UC AT users
- Serve as a resource for information about the AT, for UC campus library staff (e.g., supply informational or marketing briefs about the software, to facilitate an understanding of the software's functions)
- Conducting ongoing assessment of AT usage system-wide; we suggest that the group could reprise the survey work conducted by NGTS POT3 LT1 at a later time, in order to comparatively evaluate usage information

It is anticipated the the UC AT User Group will satisfy some of its responsibility by creating and managing an online portal that would provide ongoing support for UC AT users. This portal could include such things as best practice guidelines, information on the integration of AT record exports with existing access/discovery systems, simple tutorials, and FAQs. Appropriate links to the broader AT community and forums would also be established to support the overall use and adoption of the AT.

Composition

Membership on this group consists of archivists, librarians, curators, and other stakeholders in the UC system (primarily but not exclusively UC Libraries' staff) that use the AT to satisfy their responsibility for the description of archival materials. Each campus library will designate a member from its campus to serve on this group on a multi-year basis. That person, in addition to system-wide projects, will be responsible for coordinating AT related efforts on their campus. Such campus related efforts should not necessarily be restricted to the libraries and could extend to other campus units that are using or are interested in using the AT for the description of archival materials. Extending the concerns of the expert group beyond the campus libraries will help to establish a more uniform set of practices for using the AT and for describing archival materials throughout the UC system. Uniformity of this kind will help to decrease the cost of archival

description and improve the utility of the Online Archive of California (OAC) database of archival finding aids.

One member from the Cataloging and Metadata Common Interest Group (CAMCIG) should be appointed to serve on the group, as a formal liaison, on a multi-year basis.

Reporting Line

The UC AT User Group reports to the UC Heads of Special Collections (HOSC).

Chair

A Chair will be appointed from the group's membership, on a rotating basis by alphabetical order of the campuses. Each Chair serves a 1-year term.

Meetings

The group should normally meet on a quarterly basis, with additional business conducted by e-mail, wikis, and other communication tools.