

Catalyst Fund Proposal

Proposal Title:	WCAG 2 compliant digital archive for small library
ID:	21
Institution:	Lee College Library
Requestor:	Paul Arrigo, Director of Library and Professional Development
Budget:	\$16,475
Goal (as pulled from the application):	“Create an ADA compliant digitization software interface, a crowdsourced description creation process, and a simplified metadata methodology that will be extremely easy and efficient for small library/archive implementation.”
Description:	<p>Goal: Create an ADA compliant digitization software interface, a crowdsourced description creation process, and a simplified metadata methodology that will be extremely easy and efficient for small library/archive implementation.</p> <ul style="list-style-type: none"> • Objective 1: Persons with a visual disability will be able to understand and read the photographs being depicted in the photograph record group metadata record with screen reading software. • Objective 2: Photograph descriptions will be generated using a crowdsourcing software interface (Flickr), thus the amount of unidentified individuals and images will be reduced by 50 % . • Objective 3: Lee College and other small libraries/archives will find that the full record group metadata creation process takes exponentially less time to complete than creating a full metadata record for every single photograph. (10 minutes per record * 2500 metadata records = 25000 minutes / 60 minutes = 416 hours) vs. (10 minutes per record * 115 record group metadata records = 1150 minutes / 60 minutes = 19 hours). The 10 minutes per record time duration was determined by digitizing a single photo into TIFF and JPEG according to the digitization standards. Thus, it amounts to 416 hours vs. 19 hours with a 21 to 1 ratio. <p>Problem: Small and even medium sized libraries like Lee College’s don’t have the resources or manpower to digitize and create the metadata for a large collection of photographs efficiently without a grant. Many institutions in Texas also need to keep their webpages and digital repositories ADA compliant which can also be quite time consuming. In fact, on January 17, 2017 the United States Access Board implemented regulations making it a requirement for all federal agencies to make all of their websites compliant with WCAG 2.0 standards. Texas has had to follow Section 508 of the Rehabilitation Act since 2007. So why not come up with a solution that can deal with both issues at the same time.</p> <p>Solution: Lee College will first digitize the photos in-house according to national standards and upload them into the Flickr interface allowing the Lee College community members/disciplinary experts to provide the title and the description of the photographs. These community derived descriptions which will be used in creating ADA accessible Section 508 alternative text attributes or long descriptions. Lee College will also only create full metadata records for record groups rather than for each individual photograph. As a result, Lee College will only need to create 115 full metadata records for a collection that contains 2,500 photographs. Each of the 115 full metadata records will contain any number of photographs which fit under</p>

that provenance, with each photo given a simple or a long description in a note field, and an accompanied thumbnail link to the full image (please see the following url to get an idea of what it will look like <https://www.leecollegelibrary.com/pics/Sample % 20Full % 20Metadata % 20by % 20Record % 20Group.jpg>). The thumbnail image would have the alt or long description in the HTML for ADA purposes. The resulting larger size photograph would also have the alt text or long description terminology embedded in its HTML. Lee College will create this interface with WordPress and will either use Dublin Core or VRA Core as its metadata schema. Benefits to the LYRASIS community: Smaller libraries with few in-house resources will be able to use the Flickr crowdsourcing procedures used in this project to create metadata descriptions from their community experts. Creating full metadata records only for record groups will make it more efficient for smaller libraries to create their metadata, because they don't have to create detailed metadata for each photograph in their collection. The first methodology and the second software interface will go a long ways toward making digitization a much more efficient and achievable task for small libraries. In addition, all of the photographs, be they the thumbnail or the full image format, will use the alt text or the long description to make it ADA compliant according to WCAG 2.0 standards. This interface and methodology combination will allow LYRASIS to provide services for smaller libraries and to the ADA community. Should a small library or small archive decide to outsource the digitization the amount of work would be even more significantly reduced.

Plan of Work Timeframe:

April 2017: Donna Mohlman, a part-time librarian, will begin digitizing the photographs in two formats listed below:

- Master Files: TIF with no compression, 3000-5000 pixels across the long dimension. It will also following the following formula to determine the appropriate scanning resolution for a desired output pixel dimension: $ppi = output\ dimension / physical\ size\ of\ photo\ in\ inches$. (I.e., thus for an 8 x 10 inch photo and a desired output of 3000 pixels across the long dimension: $3000 / 10 = 300\ ppi$). For gray prints 8 bit grayscale will be utilized and for color prints 24 bits per pixel will be used. For 2500 photographs the hard disc size necessary would only be 15 Gigabytes.

- Access Files: Down-sampled (reduced in pixel dimensions) then saved as JPEGs with minimal lossey compression. Down-sampling will reduce the file size and thereby require minimal compression when saved as JPEGs. The entire photograph should be visible on the viewer's screen. To match todays screen sizes a maximum of 800 pixels across the vertical dimension will be utilized for most monitors. This process will take 5 minutes per file format. If we have 2500 photographs there will be 5000 files to create (2500 in Tiff and 2500 in JPEG). Thus the amount of time it will take to digitize all 5000 is as follows. Five minutes * 5000 files = 25,000 minutes / 60 minutes = 417 hours. 417 hours / 10 hours of work for the part time librarian per week = 41 weeks.

As they are created, the Masters (TIFFs) and Access Copies (JPEGs) will be copied and backed up to the Lee College Network Attached Storage Server. The storage server specifications include: System: Model PowerVault NX400 Processor: Intel(R)

Xeon(R) CPU E5-2403 v2 @ 1.80GHz, 1800 Mhz, 4 Core(s), 4 Logical Processor(s)
Installed Physical Memory (RAM) 8.00 GB 4 – Raid 5 3TB 7.2K RPM SAS 6Gbps 3.5in
Hard Drive.

Completion date: Based on this schedule all of the photographs will be digitized by
February 2nd 2018.

April 2017: As the files are created the JPEGs will be made available on Flickr for the
Lee College community/experts to provide descriptions.

June 2017: Will Mayer, the Distance Education and E-Resources Librarian, will
create the WordPress interface as described above.

The WordPress public interface will be loaded onto the Lee College commercial
website InMotion Hosting with the domain name of leecollegelibrary.com. This site
has unlimited storage space.

Completion date: August 31, 2017.

September 2017: The 115 full metadata records will be created by Donna Mohlman
the part-time librarian. This will take her approximately a total of 19 hours to
complete the creation of these full metadata records along with the copying and
pasting of the note field descriptions of the photographs and the ADA html coding.
This will cost an additional \$521. In addition, those photos that received
descriptions on the Flickr account will be inputted into the full metadata records as
note fields and thumbnails, with accompanied alt txt and long descriptions for ADA
accessibility. This site will not be made available to the public until all of the photos
are available and everything has been tested for quality. The last items to be
digitized will be done by February 2, 2018. The library will wait another month to
see if crowd sourcing descriptions have been made for the last photos to be
digitized in February. All of the photographs will be added to the database by March
31, 2018.

Will Mayer, who holds a TxDLA Accessibility Certification to be completed March
25th-28th, will check the WCAG 2.0 html coding that Donna inputs into the
metadata for quality control purposes.

Completion date: March 31, 2018.

April 2018: Lee College library will make the WordPress created website code and
structure open source and it will export the website's structure to a .zip file and
make the file available in an open source repository.

Budget:

- Donna Mohlman's creation of 5,000 image files from 2,500 photographs: \$11,234
 - o This process will take 5 minutes per file format. If there are 2500
photographs there will be 5000 files to create (2500 in Tiff and 2500 in
JPEG). Thus the amount of time it will take to digitize all 5000 is as follows.

	<p>Five minutes * 5000 files = 25,000 minutes / 60 minutes = 417 hours. 417 hours / 10 hours of work for the part time librarian per week = 41 weeks. 41 weeks * 27.40 hourly wage * 10 hours per week = \$11,234. The reason the library is using a part-time librarian versus a student is because the librarian will produce a much higher quality of work. The library can only afford 10 hours per week from the part-time librarian because she also provides instruction and reference desk coverage.</p> <ul style="list-style-type: none"> • Will Mayer's creation of the WordPress interface: \$4,720 <ul style="list-style-type: none"> o Mr. Mayer said it would take him a month to create this Internet interface using WordPress. Since he has other assignments he will only be able to dedicate 20 hours each month. The calculation for the total amount to pay Mr. Mayer in the form of a stipend is as follows: 160 hours (1 month) * 29.50 hourly wage = a stipend of \$4,720. Mr. Mayer has had experience with WordPress and he will also have received a TxDLA Accessibility Certification from March 25th to the 28th, 2017. • Donna Mohlman's creation of 115 full metadata records, copying and pasting of crowd sourced descriptions, and ADA html coding: \$521 <ul style="list-style-type: none"> o It will take Donna 10 minutes per record group to create each full metadata record. Thus 115 full metadata records * 10 minutes = 1150 minutes / 60 minutes = 19 hours. 19 hours * 27.40 hourly wage = \$521. Again, a part-time librarian was chosen because the attention to detail and precision work is of the utmost importance. • Total amount requested: \$16,475
<p>Comments from Field Reviewers:</p>	<ol style="list-style-type: none"> 1. Great idea! A few questions: do they have a communications plan or a way to ensure people know that images are available for crowdsourced metadata? How did they determine that the amount of unidentified individuals and images will be reduced by 50%? Is that a rough estimate? Will they provide any guidelines to the community for metadata creation? Otherwise, I imagine the info would differ widely. Love the idea and will be happy to see it tested and shared with the larger community. I don't know that a month is long enough to give people opportunity to provide metadata in the community, but I imagine that you need to create a deadline to meet the overall project timeline. 2. If funded, I would like to see the project focus more on the ADA compliance issue - this is really what is unique about this proposal. What I don't see is: (1) a full explanation of the ADA issues they are trying to address with this approach other than to meet legal requirements, (2) testing with the ADA community once the work is done (that is, has their work addressed #1), and (3) a summary of the results with the larger library, archives, museum community. If there is a re-focus on the ADA issue this would be a valuable project.