

# Community Efforts to Develop Best Practices in Digital Library Assessment: One Year of Progress

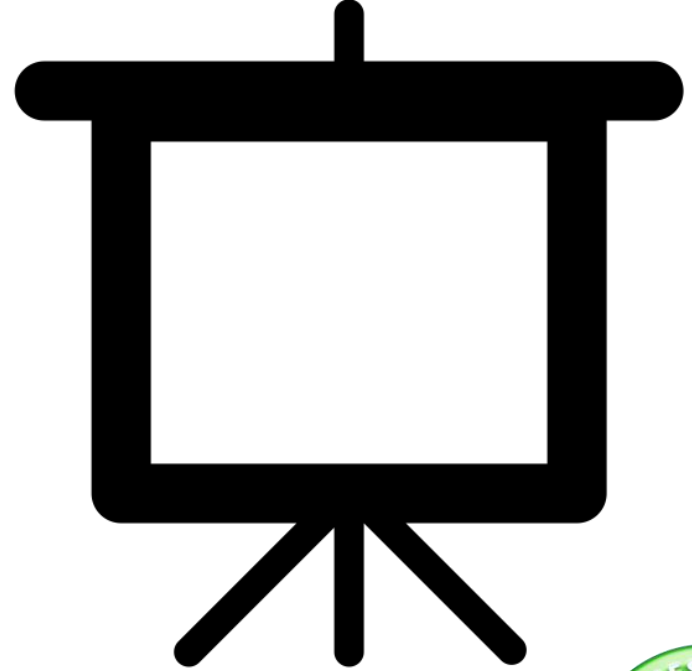
DLF Assessment Interest Group

Southeastern Library Assessment Conference  
Atlanta, Ga.  
November 17, 2015



# Presentation Outline

1. Background
2. Users Studies
3. Analytics
4. Citations
5. Cost Assessment
6. Questions



Created by ARudmann  
from Noun Project



# DLF AIG: Important Links

*You do not need to be a member of DLF to join/participate!*

DLF Assessment Interest Group Google Group and listserv:

<http://bit.ly/1G6EWQp>

DLF Assessment Interest Group wiki, including links to all the white papers and best practice documents and tools discussed today:

<http://wiki.diglib.org/Assessment>

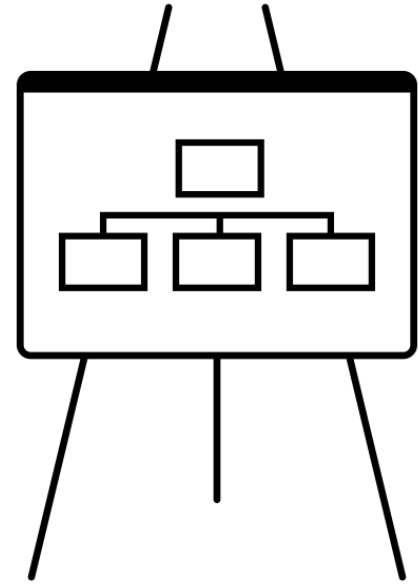
Founded in 2014, the AIG is currently co-chaired by Joyce Chapman (Duke) and Santi Thompson (University of Houston). There were 4 Working groups in 2014/15: User studies, Analytics, Citations, Costs.



# User Studies Working Group

*Evaluating users of digital libraries and their needs*

1. Goals of project
2. Methodology
3. Results and recommendations
4. Next steps



Created by José Manuel de Laá  
from Noun Project



# Goals

Guidelines/best practices for assessing facets of digital library value by analyzing user interaction with digital libraries.



Created by parkjisun  
from Noun Project



# Scope

User and Usability  
Studies



Return on Investment



Content Reuse

Created by Lemon Liu  
from Noun Project



# Examining the literature

1. What research strengths exist in the areas of usability, ROI, and reuse assessment in digital libraries?
2. What gaps exist in these areas of focus?
3. What are possible next steps for the community to address?



Created by Callum Egan  
from Noun Project



# User and Usability Studies

<b>Strengths</b>	User-centered design and assessment	Design strategies through user search behavior	
<b>Research Gaps</b>	Behavioral observations and examination of user's' task context	Over-reliance on standard testing tasks and user feedback	Lack of studies on user interactions
<b>Recommendations</b>	Users' research needs	User's role in system development	Cross-institutional collaboration to normalize usability methods



# Return on Investment

<b>Strengths</b>	Measurement of time and cost for processing	Theoretical application of ROI to library project management	
<b>Research Gaps</b>	Benefits of cost/benefit analysis	Limited corpus of cost data	No standard methodology for implementation
<b>Recommendations</b>	More studies	More data	More tools

# Content Reuse

<b>Strengths</b>	Reuse among humanities-focused digital repositories		
<b>Research Gaps</b>	Patterns of reuse	Web log analysis	Non-digital reuse
<b>Recommendations</b>	User groups -- science and social science	Digital repository interface and reuse	Assessment framework

## Next steps

Documenting best practices?

Generating assessment  
toolkit(s)?

Or...?

*Wanna get involved?*

*Contact Santi Thompson:  
sathomp3@central.uh.edu*



Created by Andrea Verzola  
from Noun Project



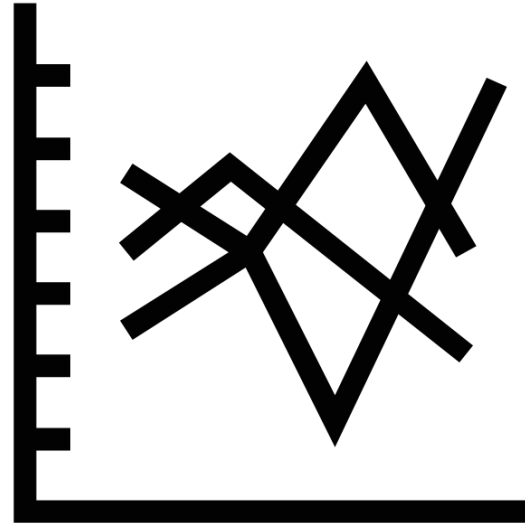
# White Paper: “Surveying the Landscape: Use and Usability Assessment of Digital Libraries”

<http://bit.ly/1KzCU6S>



# Analytics

1. Goal and scope
2. White paper methods and recommendations
3. Next steps



# Goal

*Develop guidelines and best practices for using analytics in digital libraries*

# Scope

Focus on Google Analytics

Guidelines for all experience levels



# Literature review

Literature review reveals gap in best practices

“How-to” analytics literature covers: use analytics to improve usability and discoverability, set-up google analytics (case studies), and complete transactional log analysis



# White Paper: Metrics Methodology

1. Choose metrics to recommend
2. Define each metric
3. Importance of metric
4. Bolster definition with library-centric examples

**Caveat: metrics require interpretation by local organization to be relevant and actionable.**





# Baseline Google Analytics Metrics Recommendations

## A. Content Use and Access Counts

1. Content Use and Access Counts Defined
2. Site Content Reports
3. Bounce Rate
4. Download Counts
5. Time
6. Pageviews
7. Sessions

## B. Audience Metrics

1. Location
2. Mode of Access
3. Network Domain
4. Users

## C. Navigational Metrics

1. Path Through the Site
2. Referral Traffic
3. Search Terms



# White Paper: More than Metrics

Approaching analytics: know thyself

Alternative tools and methods: consider trade-offs between tools and methods

Going beyond the baseline: customization and platform specific considerations



# Next steps

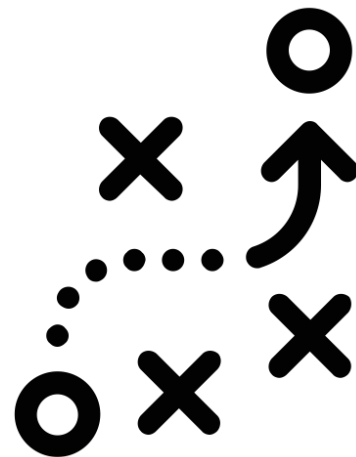
Options options options

More platform specific metrics?

More examples?

Share metrics?

Discontinue the analytics group?



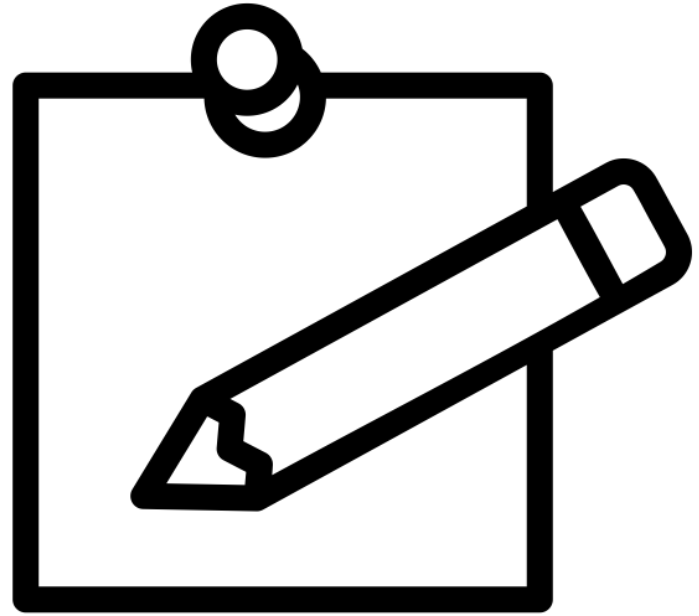
# White Paper: “Best Practices For Google Analytics in Digital Libraries”

<http://bit.ly/1Pp11NF>



# Citations Working Group

1. Background discussion
2. Methodology
3. Recommendations
4. Next steps



Created by Johan H. W. Basberg  
from Noun Project



# Citations Working Group

## DLF AIG Charge

1. What should a citation consist of?
2. How can we best support appropriate citations?
3. To what extent do common citation formats support this?
4. What are the limitations of current digital library software systems for displaying citation information?
5. What are best practices for displaying citation information for reference manager software capture?



# Citations Working Group

Draft citation standards, based on what can and can't be incorporated into APA, Chicago, and MLA, that incorporate the necessary elements for digitized special collections and institutional repository content



# Citations Working Group

## Methodology

Data set citations

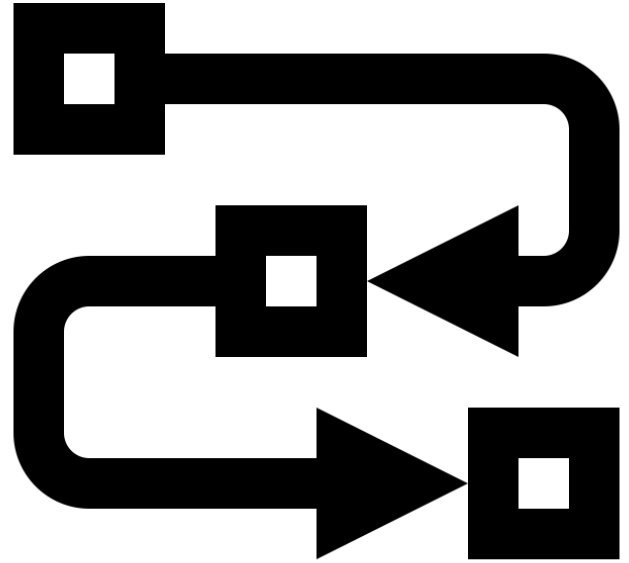
Archival/manuscript citations

Web/electronic document citations

Institutional citation recommendations

## Citation Styles

APA, Chicago (Notes and Bibliography), MLA



Created by Rflor  
from Noun Project





# Citations Working Group

## Recommendations

Include:

- Item name or title
- Collection name
- Repository information (physical and/or digital)
- Unique identifier (Digital Object Identifiers (DOI), Handle, Persistent Uniform Resource Locators (PURL), citable URL, etc.)



Created by Gregor Črešnar  
from Noun Project



# Citations Working Group

Next Steps for these guidelines:

Publication

Adoption

Advocacy



Created by John Winowiecki  
from Noun Project



# Citations Working Group

## Next steps for citations working group:

1. What are the limitations of current digital library software systems for displaying citation information?
2. What are best practices for displaying citation information for reference manager software capture?
3. More? Do we as a group start creating scripts/plugins so that the most common DL & IR platforms can automatically generate statements for preferred citations? Do we need to do a bibliometric study of DL items in scholarly literature? Etc.



White Paper: “Guidelines for citing library-hosted,  
unique digital assets”

<http://bit.ly/1MNJ3Ci>

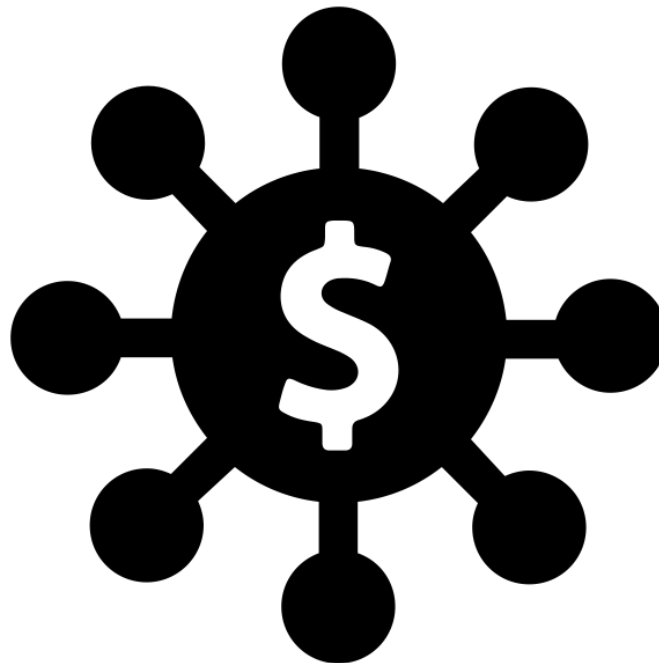


# Cost Assessment Working Group

Goals

Process & products

Call for data



Created by Gregor Črešnar  
from Noun Project



# Cost Assessment Working Group

## Goals

Process & products

Call for data

*To aggregate and make freely available a large set of time/cost data on the performance of various tasks involved in the digitization process, in order to assist organizations in digitization project planning and benchmarking.*



# Cost Assessment Working Group



## Library Digitization Cost Calculator

calculator

notes on data

feedback / data submission

This calculator aggregates available data on the cost and time it takes to perform various activities associated with library digitization. The calculator provides estimates not accuracy, as each institution is different. The tool provides average time and cost information from other institutions who have donated their data to this project. Data is currently limited. This calculator was built by Joyce Chapman, please contact [joyce.chapman@duke.edu](mailto:joyce.chapman@duke.edu) for more information, or use the [feedback form](#).

### Types of scanner

Extent (linear feet)

Hourly pay of student workers

Student benefits as % of pay

Annual salary of staff member

Staff benefits as % of pay

Metadata creation

Average number of scans per item

Include fastener removal?

Include condition review?

Include intellectual property review?

Include time to sort materials into items?

Include flattening/supporting materials pre-scanning?

Include time for fragile item handling?

Include time for disbinding and rebinding items?

Select

0

Assumed: 1 lf = 1,200 scans

0

E.g., 10.25

0

E.g., 20.3

0

E.g., 45000

0

Select

To calculate metadata creation time (by item, not scan)

Yes Calculations will be for removing fasteners from 100% of items

Yes Flagging/pulling materials to be sent to conservation department.

Yes Calculations will be for review of 100% of items

Yes E.g., locate pp. 1-5 of a discrete document and place them together

Yes This will be calculated as if it were done for 100% of items

Yes This will be calculated as if 100% of items are fragile

Yes This will be calculated as if 100% of items need disbinding

Calculate

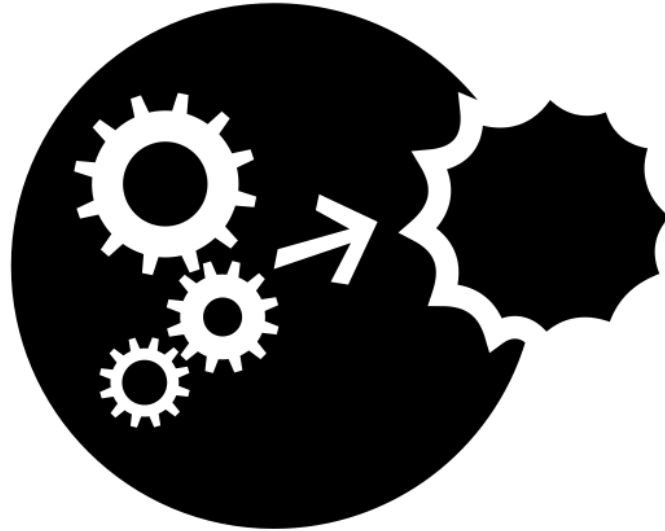


# Cost Assessment Working Group

Goals

**Process & products**

Call for data



Created by Takao Umehara  
from Noun Project





# Cost Assessment Working Group

Performed a review of existing literature (published and unpublished) in the areas of

- Collection of time and cost data for digitization
- Existing best practices in quality control and metadata creation
- Found fewer than 20 resources!



# Cost Assessment Working Group

## Guidelines and definitions (1 of 2)

### Preparation of original materials

Condition review

Disbinding

Fastener removal

Flattening

Rights review

Sorting materials into items

Supporting

Unique identifier assignment

### Image capture

Film or transparency scanner

Flatbed scanner

Manual DSLR camera

Medium format camera

Overhead scanner

Sheet feeding scanner



# Cost Assessment Working Group

## Guidelines and definitions (2 of 2)

Quality control

Level 1, 2 or 3

Descriptive metadata creation

Level 1, 2 or 3

Post-processing

Alignment/rotation

Background removal

Clean up / dust removal

Color correction and tonal  
adjustments

Cropping images

Stitching

Post-preparation

De-sorting

Re-binding

Re-fastening



# Digitization Cost Calculator

calculator	notes on data	feedback/data submission
------------	---------------	--------------------------

Extent to digitize

**People performing the work**

Salaried 1

Hourly 1

.....

**Preparation of original materials**

Condition review

Disbinding

Fastener removal

.....

**Image capture**

Capture device

**Quality control**

Level 1

Level 2

Level 3

**Post-processing**

Alignment/rotation

Background removal

.....

**Local Fields**

Local field 1

Local field 2

# Digitization Cost Calculator

calculator	notes on data	feedback/data submission
------------	---------------	--------------------------

**Results**

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officina deserunt mollit anim id est

Activity	Time (hours)	Cost (Salary 1)	Cost (Hourly 1)...
Fastener removal	10.3	N/A	\$\$\$
Image capture	34.7	\$\$\$	\$\$\$
Metadata creation	62	\$\$\$	N/A
.....	.....	.....	.....
<b>Total</b>	<b>105</b>	<b>\$\$\$</b>	<b>\$\$\$</b>



# Cost Assessment Working Group

## Call for data submissions

[Call for data submissions](#) (a document that explains what we're doing and tells you how to contribute your data)

[Data submission form](#) (the actual online form for submitting data),  
or just copy this link <http://bit.ly/1LV9oxl>

*All data submitted to this project will be publicly available, both via aggregate calculations made by the calculator, and by institution on a separate reference page of the calculator's website.*



# Cost Assessment Working Group

## Call for data submissions

What amount of data do I have to submit?

What format do I submit the data in?

How does the calculator work / how will my submitted data be used?



# Questions?

Contact: [joyce.chapman@duke.edu](mailto:joyce.chapman@duke.edu)

Join the Google Group & listserv: <http://bit.ly/1G6EWQp>

Download documents and follow progress on our wiki:  
<http://wiki.diglib.org/Assessment>

