RAMP:

Repository Analytics and Metrics Portal

Accurately measuring use of institutional repositories

Kenning Arlitsch, Dean **@kenning_msu**
Patrick OBrien, Semantic Web Research Director **@sempob**

ARL Assessment Forum
Atlanta, GA, January 20, 2017

Agenda

1. A New Reporting Model
2. Accuracy of Analytics Reporting Tools
   - Page Tagging
   - Log file
3. RAMP
   - New Prototype Web Service for IR reporting

A New Reporting Model

Page Type	Definition	Examples
Citable Content Downloads	Non-HTML scholarly content that may be formally cited in the research process	• Publication (.pdf) • Presentation (.ppt) • Data Sets (.csv)
Item Summary	HTML pages to help user decide to download the full publication	• Title & Abstract • Item Metadata
Ancillary	HTML pages that provide general information or navigation	• Search Results • Browse by Author • Statistics

ScholarWorks

Open Access Scholarship at Montana State University

Search articles, professional papers, theses, dissertations

Business, Economics & Management
Chemical & Material Sciences
Engineering & Computer Science
Health & Medical Sciences
Humanities, Literature & Arts
Life Sciences & Earth Sciences
Physics & Mathematics
Social Sciences

DISCOVER

Author

Arlitsch, Kenning (4)

Mixter, Jeff (4)

OBrien, Patrick (4)

Sterman, Leila (3)

Borda, Susan (2)

Clark, Jason A. (2)

Wheeler, Jonathan (2)

Young, Scott W.H. (2)

Banner, Katie (1)

Border, J. Kent. (1)

... View More

Date Issued

2000 - 2017 (15)

Search

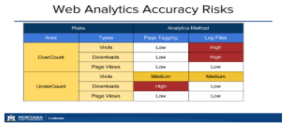
All of ScholarWork

Undercounting

Go

Show Advanced Filters

Now showing items 1-10 of 23



Data set supporting study on **Undercounting** File Downloads from Institutional Repositories [dataset]

OBrien, Patrick; Arlitsch, Kenning; Sterman, Leila; Mixter, Jeff; Wheeler, Jonathan; Borda, Susan (Montana State University ScholarWorks, 2016-07)

Search term found in abstract:...This dataset supports the study published as "**Undercounting** File Downloads from IR". The following items are included: 1. gaEvent.zip = PDF exports of Google Analytics Events reports for each IR. 2. gaItemSummaryPageViews.zip = PDF exports...

Undercounting File Downloads from Institutional Repositories

OBrien, Patrick; Arlitsch, Kenning; Sterman, Leila; Mixter, Jeff; Wheeler, Jonathan; Borda, Susan (Emerald, 2016-10)

A primary impact metric for institutional repositories (IR) is the number of file downloads, which are commonly measured through third-party web analytics software. Google Analytics, a free service used by most academic ...

(Search term found in fulltext file)

Undercounting File Downloads from Institutional Repositories

View/Open
Preprint (833.8Kb)

A primary impact metric for institutional repositories (IR) is the number of file downloads, which are commonly measured through third-party web analytics software. Google Analytics, a free service used by most academic libraries, relies on HTML page tagging to log visitor activity on Google's servers. However, web aggregators such as Google Scholar link directly to high value content (usually PDF files), bypassing the HTML page and failing to register these direct access events. This paper presents evidence of a study of four institutions demonstrating that the majority of IR activity is not counted by page tagging web analytics software, and proposes a practical solution for significantly improving the reporting relevancy and accuracy of IR performance metrics using Google Analytics.

Date
2016-10

Author
OBrien, Patrick
Arlitsch, Kenning
Sterman, Leila
Mixter, Jeff
Wheeler, Jonathan
Borda, Susan

URI
http://scholarworks.montana.edu/xmlui/handle/1/9943

Related Material/Data
http://scholarworks.montana.edu/xmlui/handle/1/9939

Collections
Scholarly Work - Library

Undercounting File Downloads from Institutional Repositories

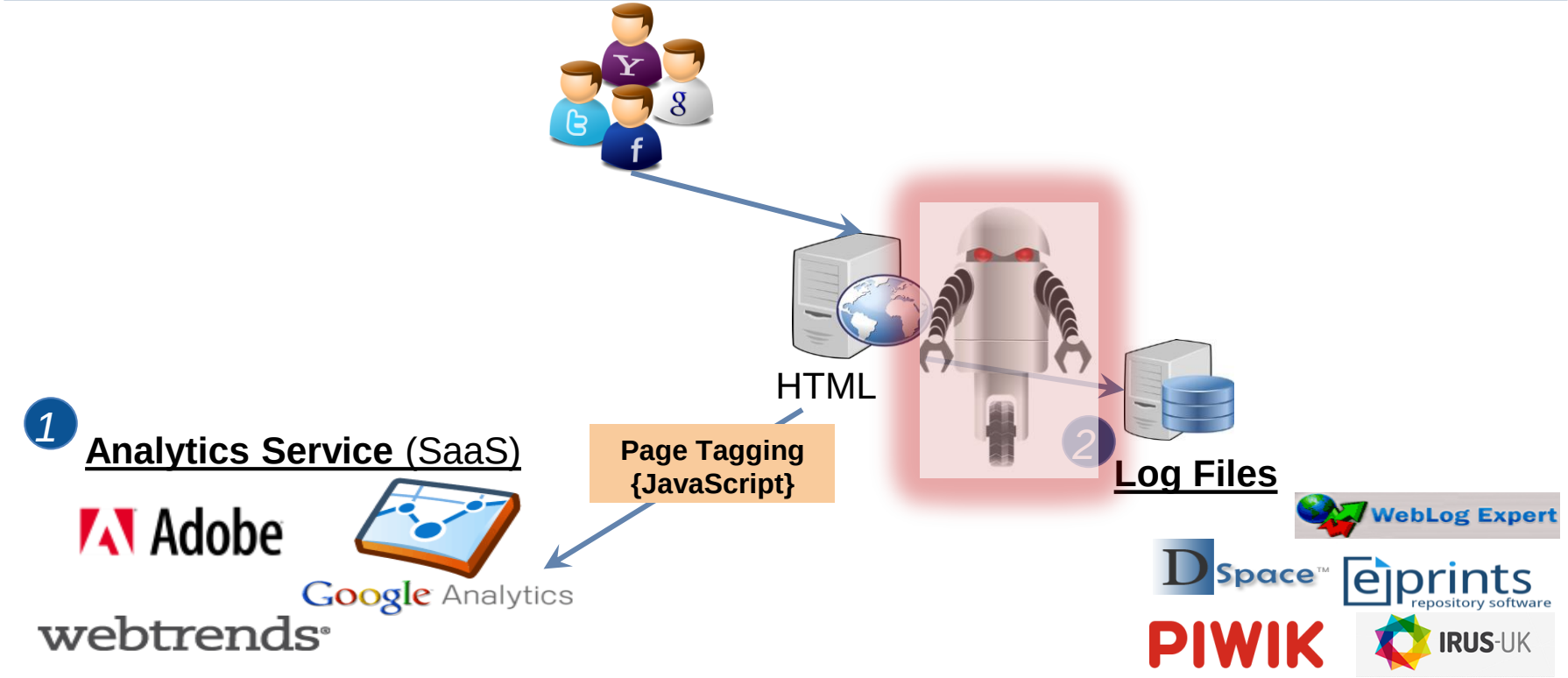
Patrick OBrien, Kenning Arlitsch, Leila Sterman, Jeff Mixter, Jonathan Wheeler, and Susan Borda

This is a preprint of an article that will appear in the Journal of Library Administration in October 2016.

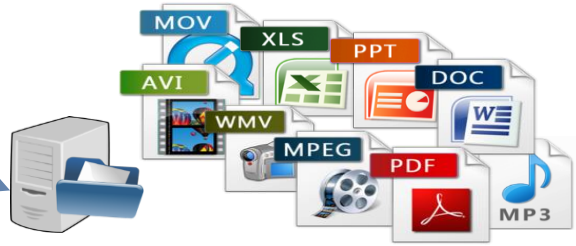
Journal of Library Administration

Accuracy of Analytics Reporting Tools

Two Classes of Web Analytics



Page Tagging *does not* track non-HTML CCD



Non-HTML

HTML

Page Tagging

Google

Introducing the" Getting Found" Web Analytics Cookbook for Monitoring

Scholar

Articles

Case law

My library

Any time

Since 2016

Since 2015

Since 2012

Custom range...

Introducing the" Getting Found" Web Analytics Cookbook for Monitoring Search Engine Optimization of Digital Repositories

[PDF] montana.edu

K Arlitsch, P OBrien - Qualitative and Quantitative ... 2015 - scholarworks.montana.edu

A new toolkit that helps libraries establish monitoring of the search engine optimizat of the products of research funded by the

Related articles All 5 versions Cite Sa

Showing the best result for this search. Se

About Google Sch

Kenning Arlitsch

April 1 ·

SEO Cookbook article finally got published.

Introducing the "Getting Found" Web Analytics Cookbook for Monitoring Search Engine Optimization...

SCHOLARWORKS.MONTANA.EDU

Like Comment Share

Kenning Arlitsch

@kenning_msu

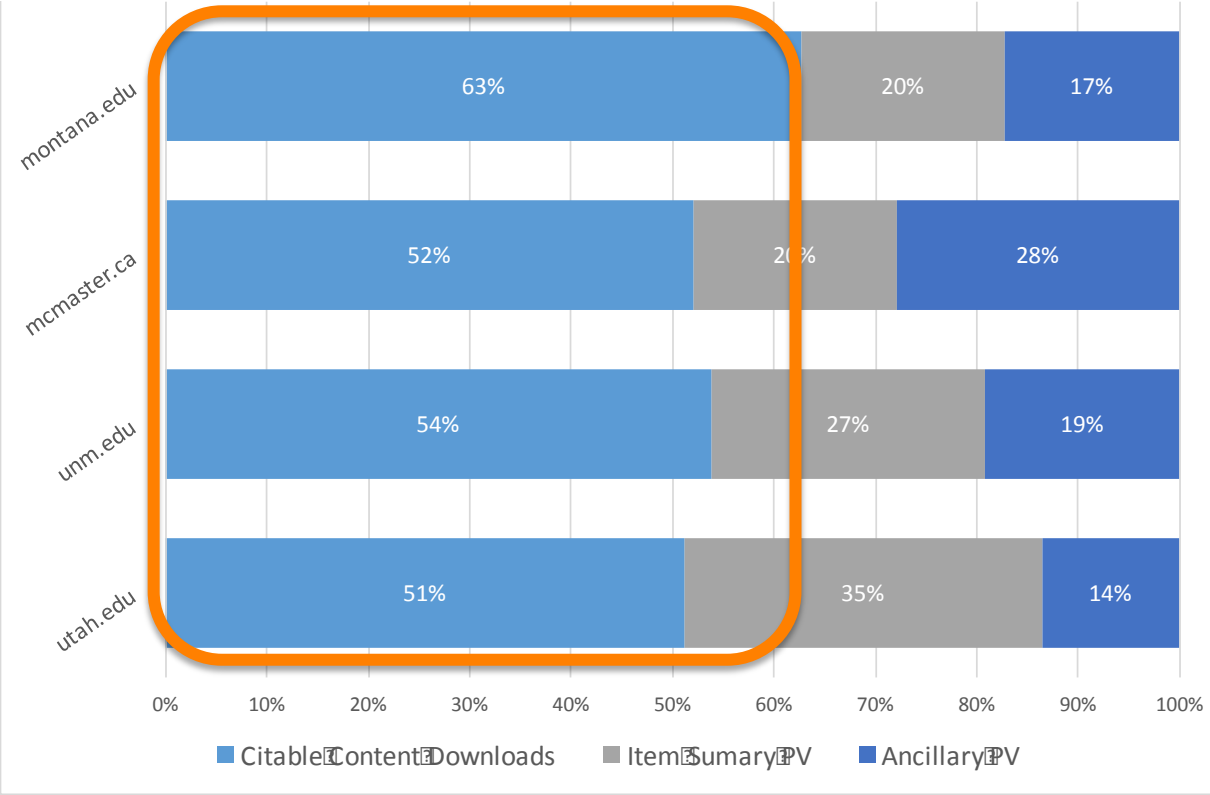
SEO Cookbook article co-authored with Patrick OBrien @sempob finally published - with scholarworks.montana.edu/xmlui/handle/1…

Google Analytics use in academic libraries

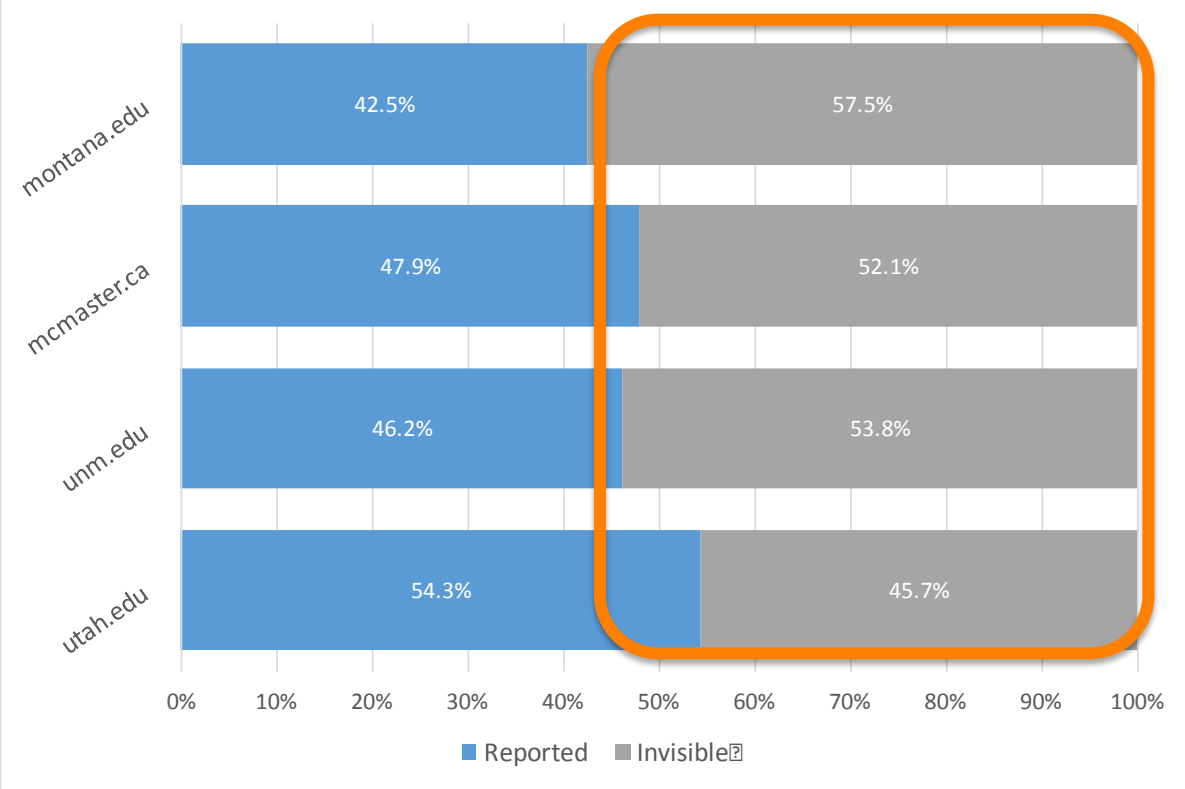
- Tested 279 academic library websites
  - ARL
  - DLF
  - OCLC-RLP

- 90% US libraries contained Google tracking code

Most IR activity is Citable Content Downloads



Most IR Activity Unreported by Google Analytics



Web Analytics Accuracy Risks

Risks		Analytics Method	
Area	Types	Page Tagging	Log Files
OverCount	Visits	Low	High
	Downloads	Low	High
	Page Views	Low	Low
UnderCount	Visits	Medium	Medium
	Downloads	High	Low
	Page Views	Low	Low

Data set: Jan 5 - May 17, 2016 (n = 134 days)

Study Participant	IR	Platform	URL
Montana State University	ScholarWorks	DSpace	scholarworks.montana.edu
McMaster University	MacSphere	DSpace	macsphere.mcmaster.ca
University of New Mexico	LoboVault	DSpace	repository.unm.edu
University of Utah	USpace	CONTENTdm	uspace.utah.edu

Page Tagging *does not* track non-HTML Citable Content Downloads

Email

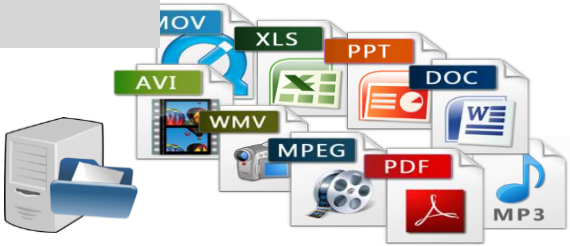
Google Scholar

Wikimedia Commons

Google SEARCH CONSOLE

Search Analytics

Does!



Non-HTML

Typical Page Tagging (GA) Data Types

Page Type	Analytics	
	Pages	Events
Citable Content Downloads	-	✓
Item Summary	✓	-
Ancillary	✓	-

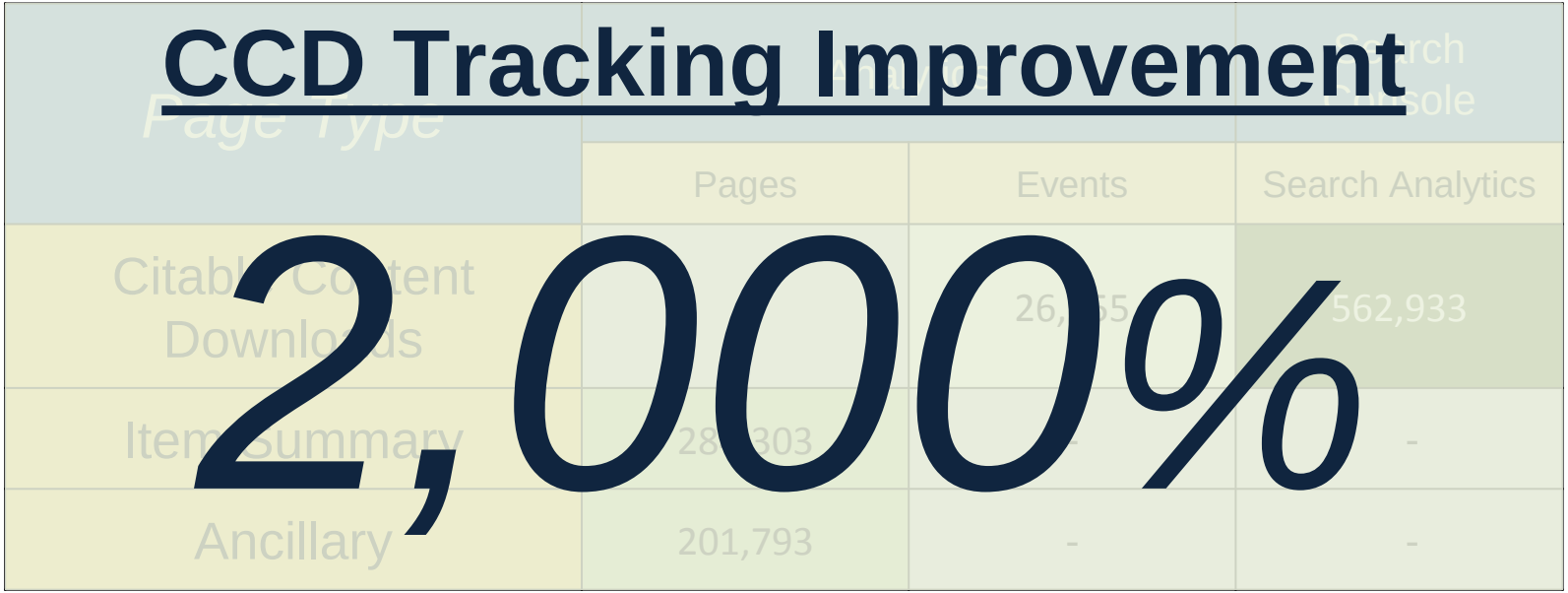
Montana Method Includes Google Search Console

Page Type	Analytics		Search Console
	Pages	Events	Search Analytics
Citable Content Downloads	-		
Item Summary		-	-
Ancillary		-	-

GA Ancillary Page Views and Item Summary Page Views vs CCD

IR	Item Sumary PV	Ancillary PV	Total Google Analytics HTML PV	Download Events	Citable Content Downloads
scholarworks.montana.edu	26,735	23,350	50,085	7,129	77,380
macsphere.mcmaster.ca	51,150	71,585	122,735	n/a	133,342
repository.unm.edu	83,491	59,289	142,780	n/a	166,320
content.lib.utah.edu	122,927	47,569	170,496	19,226	159,536

All four IR using Google Analytics



Montana Method Challenges

- Missing non-Google direct link CCD events
  - Yahoo
  - Bing
  - Email
  - FB
- GSC limits time and access
  - Moving 90-day window
  - Granular data = programing skills to access API

RAMP: Repository Analytics & Metrics Portal

RAMP - Repository Analytics and Metrics Portal

- Cloud Web service
- Currently accessible and free
  - During grant period (through November 2017)
- No training or configuration required
- Consistent method and terminology
- Benchmarking across time and organization
- Request access - mixterj@oclc.org

Repository Analytics & Metrics Portal

Search Here

Montana State University

This is for text about the sites that have Google Search API hooked up

University of New Mexico

This is for text about the sites that have Google Search API hooked up

McMaster University

This is for text about the sites that have Google Search API hooked up

About the RAMP Portal

Montana State University, the Association of Research Libraries, the University of New Mexico, and OCLC Research have joined as partners to examine the difficulties that libraries face in producing accurate reports on the use of their digital repositories.

Repository Analytics & Metrics Portal

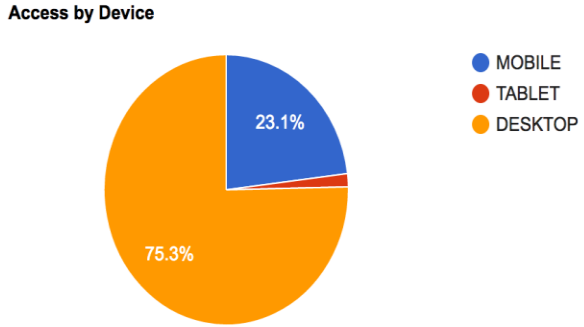
Montana State University

Access Data

Selected Date: Wednesday, January 11, 2017

	January 2017						
	Mon	Tue	Wed	Thu	Fri	Sat	Sun
52	26	27	28	29	30	31	01
1	02	03	04	05	06	07	08
2	09	10	11	12	13	14	15
3	16	17	18	19	20	21	22
4	23	24	25	26	27	28	29
5	30	31	01	02	03	04	05

View Device Access Chart | Download Access Data



Publications

Published:

Patrick OBrien, Kenning Arlitsch, Leila Sterman, Jeff Mixter, Jonathan Wheeler, and Susan Borda. “Undercounting File Downloads from Institutional Repositories,” *Journal of Library Administration*, vol. 56, no. 7, 2016

Forthcoming:

Patrick OBrien, Kenning Arlitsch, Leila Sterman, Jeff Mixter, Jonathan Wheeler, and Susan Borda. “RAMP: Repository Analytics and Metrics Portal: A Prototype Web Service that Accurately Counts Item Downloads from Institutional Repositories,” Accepted by *Library Hi Tech*, expected early 2017

Proposal funded by IMLS:

”Measuring Up: Assessing Accuracy of Reported Use and Impact of Digital Repositories” - scholarworks.montana.edu/xmlui/handle/1/8924

Undercounting Research Team

- Montana State University
  - Kenning Arlitsch, Dean **@kenning_msu**
  - Patrick OBrien, Semantic Web Research Director **@sempob**
  - Leila Sterman, Scholarly Communication Librarian **@calamityleila**
  - Susan Borda, Digital Technologies Librarian **@mutanthumb**
- OCLC Research
  - Jeff Mixter, Software Engineer **@jeffmixter**
- University of New Mexico
  - Jonathan Wheeler, Data Curation Librarian