

RAMP:

# Repository Analytics and Metrics Portal

Accurately measuring use of institutional repositories



Kenning Arlitsch, Dean @kenning\_msu  
Patrick OBrien, Semantic Web Research Director @sempob

OCLC RLP Meeting  
London, UK, March 8, 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	<ul style="list-style-type: none"><li>• Publication (.pdf)</li><li>• Presentation (.ppt)</li><li>• Data Sets (.csv)</li></ul>
Item Summary	HTML pages to help user decide to download the full publication	<ul style="list-style-type: none"><li>• Title &amp; Abstract</li><li>• Item Metadata</li></ul>
Ancillary	HTML pages that provide general information or navigation	<ul style="list-style-type: none"><li>• Search Results</li><li>• Browse by Author</li><li>• Statistics</li></ul>

# 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)

Mixer, Jeff (4)

O'Brien, 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

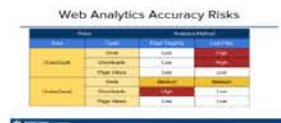
All of ScholarWork

### Undercounting

Go

[Show Advanced Filters](#)

Now showing items 1-10 of 23



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

O'Brien, Patrick; Arlitsch, Kenning; Sterman, Leila; Mixer, 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. galtemSummaryPageViews.zip = PDF exports...



## Undercounting File Downloads from Institutional Repositories

O'Brien, 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)


☒ Search ScholarWorks

☐ This Collection

**BROWSE**

All of ScholarWorks

Communities &amp; Collections

By Issue Date

Authors

Titles

Departments

Item Type

This Collection

By Issue Date

Authors

Titles

# Undercounting File Downloads from Institutional Repositories


**View/Open**
 Preprint (833.8Kb)

**Date**

2016-10

**Author**

O'Brien, Patrick  
Arlitsch, Kenning  
Sterman, Leila  
Mixter, Jeff  
Wheeler, Jonathan  
Borda, Susan

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.

**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  
Mixer, Jonathan Wheeler, and Susan Borda

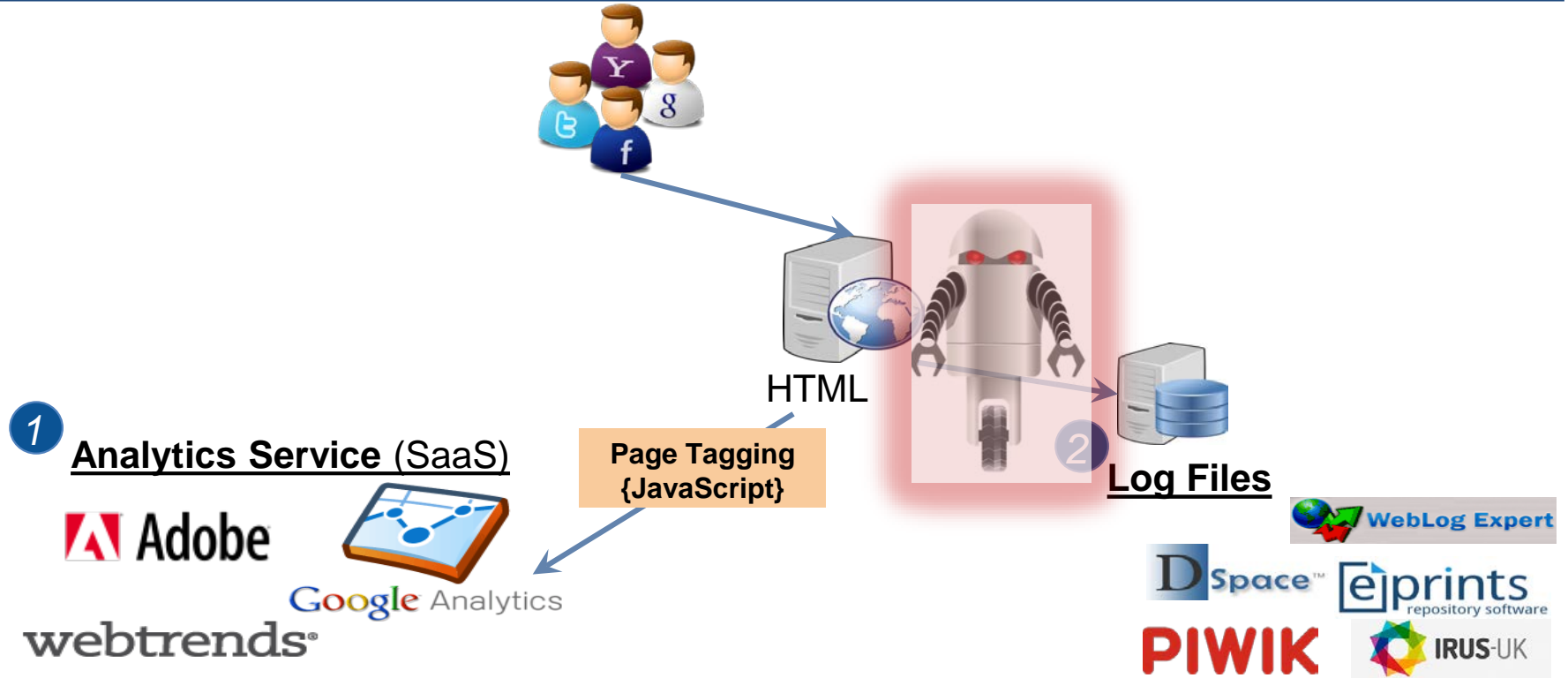
This is a preprint of an article that will appear in the Journal of Library Administration in  
October 2016.

[Journal of Library Administration](#)

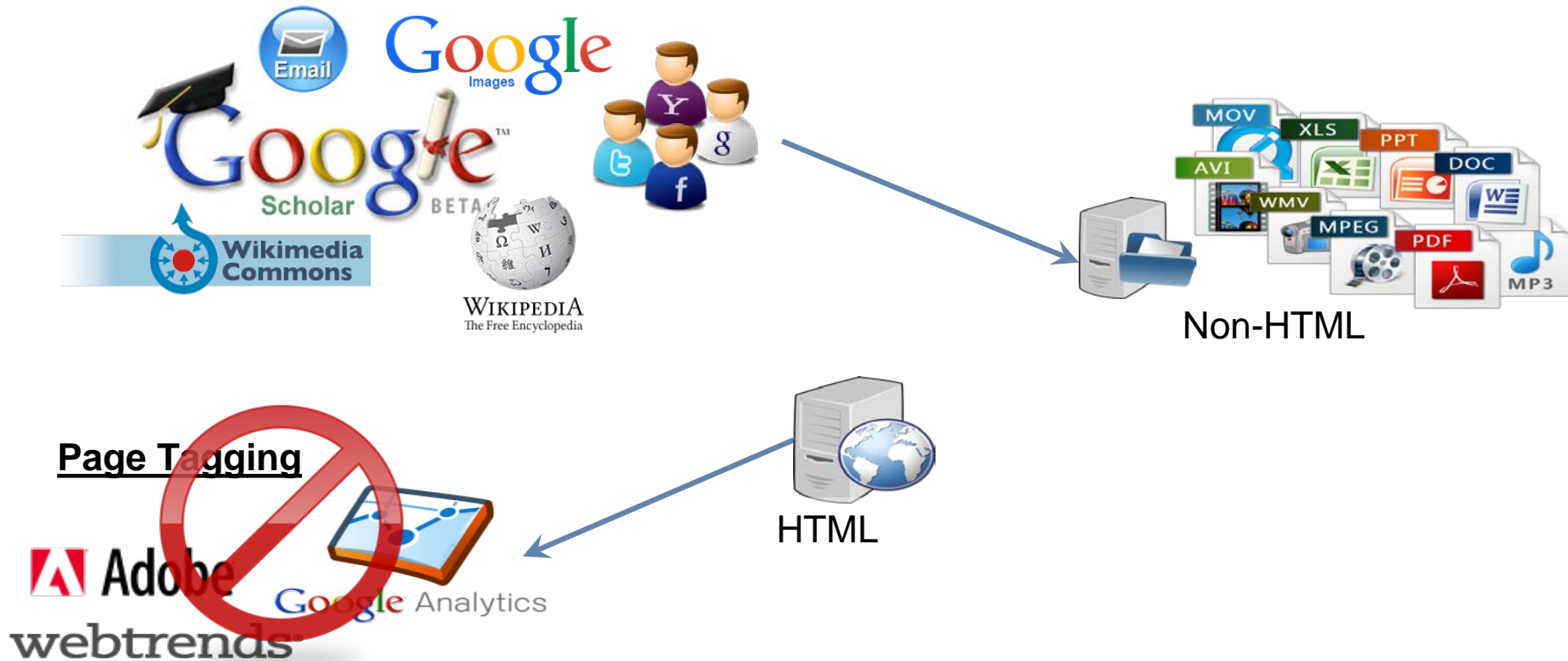
---

# Accuracy of Analytics Reporting Tools

# Two Classes of Web Analytics



# Page Tagging does not track non-HTML CCD



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

[K Arlitsch](#), [P OBrien](#) - Qualitative and Quantitative - 2015 - scholarworks.montana.edu

A new toolkit that helps libraries establish monitoring of the search engine optimization of the products of research funded by the  
[Related articles](#) [All 5 versions](#) [Cite](#) [S](#)

**Kenning Arlitsch**

April 1 ·

Showing the best result for this search. [S](#)

[About Google Scholar](#)

SEO Cookbook article finally got published.

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

[SCHOLARWORKS.MONTANA.EDU](#)

**Kenning Arlitsch**

@kenning\_msu



Like



Comment



Share

SEO Cookbook article co-authored with P. OBrien

OBrien [@sempob](#) finally published - with

[scholarworks.montana.edu/xmlui/handle/1...](#)

[\[PDF\] montana.edu](#)

# Google Analytics use in academic libraries

---

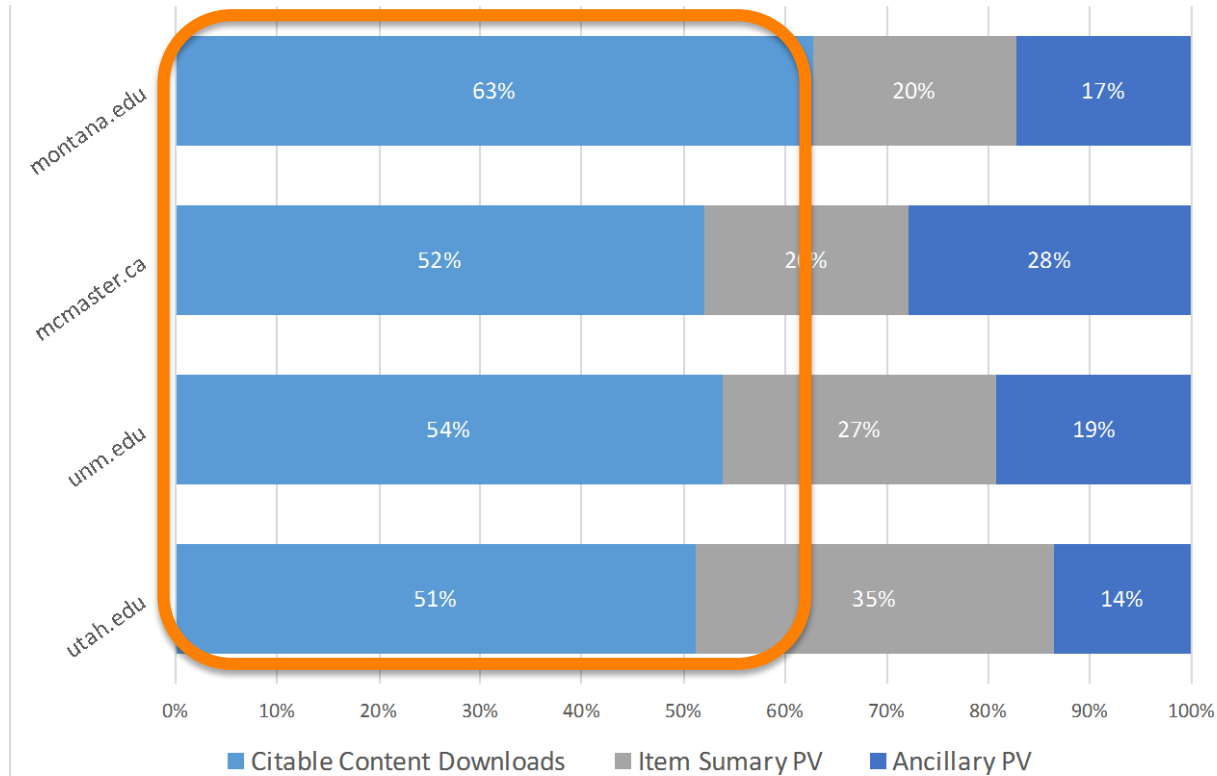
- Tested 279 academic library websites
  - ARL
  - DLF
  - OCLC-RLP
- 90% US libraries contained Google tracking code

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

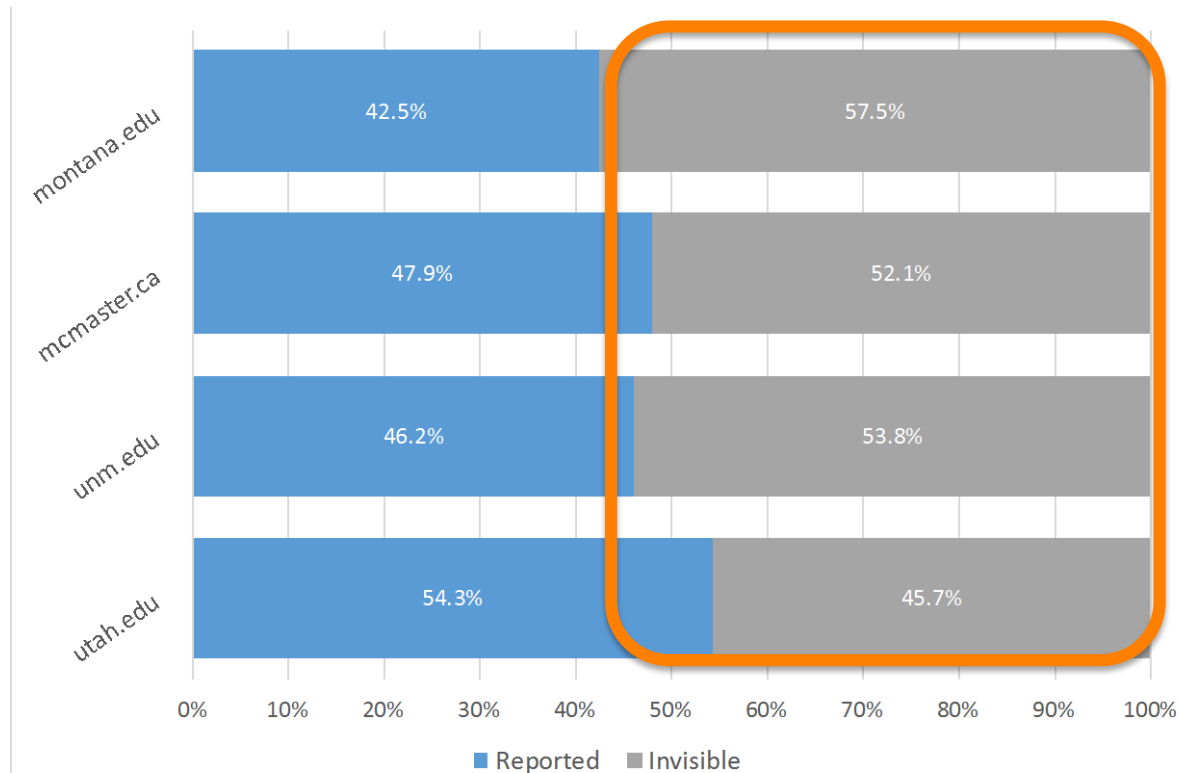
---

Study Participant	IR	Platform	URL
Montana State University	ScholarWorks	DSpace	<a href="http://scholarworks.montana.edu">scholarworks.montana.edu</a>
McMaster University	MacSphere	DSpace	<a href="http://macsphere.mcmaster.ca">macsphere.mcmaster.ca</a>
University of New Mexico	LoboVault	DSpace	<a href="http://repository.unm.edu">repository.unm.edu</a>
University of Utah	USpace	CONTENTdm	<a href="http://uspace.utah.edu">uspace.utah.edu</a>

# Most IR activity is Citable Content Downloads



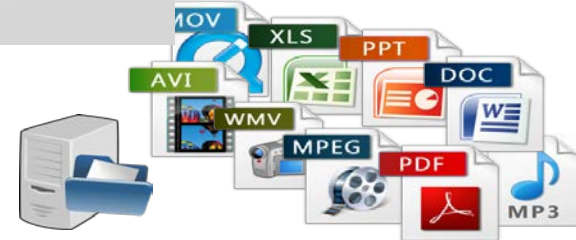
# Most IR Activity Unreported by Google Analytics



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



Does!



Non-HTML

# Typical Page Tagging (GA) Data Types

---

<i>Page Type</i>	Analytics	
	Pages	Events
Citable Content Downloads	-	✓
Item Summary	✓	-
Ancillary	✓	-

# Montana Method Includes Google Search Console

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

# All four IR using Google Analytics

---

## CCD Tracking Improvement

2,000%

Page Type	Pages	Events	Search Analytics
Citable Content Downloads	-	26,315	562,933
Item Summary	284,003	-	-
Ancillary	201,793	-	-

# 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 = programming skills to access API

---

# RAMP: Repository Analytics & Metrics Portal


# RAMP - Repository Analytics and Metrics Portal


---

- Cloud Web service
- Currently accessible and free
- No training or configuration required
- Consistent method and terminology
- Benchmarking across time and organization
- Request access - [mixterj@oclc.org](mailto:mixterj@oclc.org)

# RAMP institutions as of March 3, 2017


Currently tracking over 86,000 digital items from 7 IR and capturing over 10,000 CCD per day that were previously invisible.

**Repository Analytics & Metrics Portal**




### Montana State University

ScholarWorks is an open access institutional repository for the capture of the intellectual work of Montana State University.




### University of New Mexico

LoboVault is UNM's Institutional Repository. It hosts scholarly publications from UNM faculty, graduate student theses and dissertations, UNM administrative records, and more.




### McMaster University

MacSphere is McMaster University's Institutional Repository (IR). MacSphere aims to bring together all of a University's research under one umbrella, in order to preserve and provide access to that research. The research and scholarly output included in MacSphere has been selected and deposited by the individual university departments and centres on campus.




### Maryland MDSOAR

MD-SOAR is a shared digital repository platform for eleven colleges and universities in Maryland. It is jointly governed by all participating libraries, who have agreed to share policies and practices that are necessary and appropriate for the shared platform.




### Maryland DRUM

The Digital Repository at the University of Maryland (DRUM) collects, preserves, and provides public access to the scholarly output of the university. Faculty and researchers can upload research products for rapid dissemination, global visibility and impact, and long-term preservation.




### OAKTrust digital repository at Texas A&M

The OAKTrust digital repository at Texas A&M is a digital service that collects, preserves, and distributes the scholarly output of the University. The repository facilitates open access scholarly communication while preserving the scholarly legacy of the Texas A&M community.



### Digital Collections of Colorado

The Digital Collections of Colorado is a digital service that collects, preserves and distributes digital material provided by a group of Colorado institutions for digital preservation and scholarly communication.








### Michigan DeepBlue

Deep Blue is the University of Michigan's institutional repository service. It preserves and provides access to the research and creative work done by our faculty, staff, and students.

### 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.





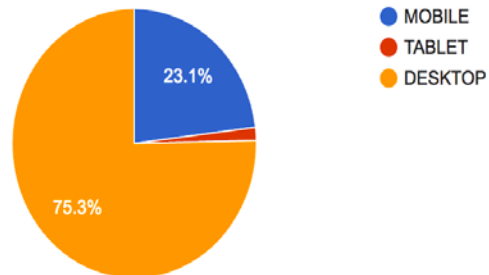
### Access Data

Selected Date: Wednesday, January 11, 2017

[View Device Access Chart](#)[Download Access Data](#)

	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

Access by Device





# Repository Analytics & Metrics Portal



Montana State University

## Access Data

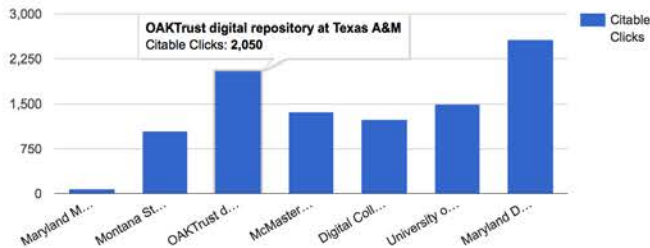
<	February 2017							>
	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
5	30	31	01	02	03	04	05	
6	06	07	08	09	10	11	12	
7	13	14	15	16	17	18	19	
8	20	21	22	23	24	25	26	
9	27	28	01	02	03	04	05	
10	06	07	08	09	10	11	12	

View Device Access Chart

View Click Comparison Chart

Download Access Data

Citable Item Click Comparison



## Settings

### Sites to Analyze

http://scholarworks.montana.edu/  
https://scholarworks.montana.edu/

Edit

Currently Logging: Yes

Stop logging

# 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. "RAMP: Repository Analytics and Metrics Portal: A Prototype Web Service that Accurately Counts Item Downloads from Institutional Repositories," *Library Hi Tech*, v35n1, March 2017

## **Proposal funded by IMLS:**

"Measuring Up: Assessing Accuracy of Reported Use and Impact of Digital Repositories" - [scholarworks.montana.edu/xmlui/handle/1/8924](https://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

## ❖ OCLC Research

- Jeff Mixter, Software Engineer @jeffmixter

## ❖ University of New Mexico

- Jonathan Wheeler, Data Curation Librarian