

# Measuring Reuse of Digital Objects: Preliminary Findings from the IMLS-funded project

## Extended Abstract

Elizabeth Joan Kelly  
Loyola University New Orleans  
USA  
ejkelly@loyno.edu

Caroline Muglia  
University of Southern California  
USA  
muglia@usc.edu

Genya O'Gara  
Virtual Library of Virginia  
USA  
gogara@gmu.edu

Ayla Stein  
UIUC  
USA  
astein@illinois.edu

Santi Thompson  
University of Houston  
USA  
sathompson3@uh.edu

Liz Woolcott  
Utah State University  
USA  
liz.woolcott@usu.edu

## ABSTRACT

The first step of the Developing a Framework for Measuring Reuse of Digital Objects project involved a survey identifying how cultural heritage organizations currently assess digital library reuse, barriers to assessing reuse, and community priorities for potential solutions and next steps. This poster offers initial analysis of the survey results.

## CCS CONCEPTS

• **Information systems** → **Digital libraries and archives**;

## KEYWORDS

digital library, reuse, e-Resources, lifecycle management, usage, collection analysis, scholarly communication

## ACM Reference Format:

Elizabeth Joan Kelly, Caroline Muglia, Genya O'Gara, Ayla Stein, Santi Thompson, and Liz Woolcott. 2018. Measuring Reuse of Digital Objects: Preliminary Findings from the IMLS-funded project: Extended Abstract. In *Proceedings of Joint Conference on Digital Libraries (JCDL '18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3197026.3203871>

## 1 INTRODUCTION

Content reuse, or how often and in what ways digital library materials are employed and repurposed, is a key indicator of the impact of digital libraries (as defined by Krystyna Matusiak)[1].

A literature review and whitepaper authored by the Digital Library Federation Assessment Interest Group (DLF-AIG) in 2015 found that traditional library analytics focus primarily on access and download statistics, which do not show how users utilize or transform unique materials from digital collections [2]. This lack of distinction, combined with nonstandard assessment approaches, makes it difficult to develop user-responsive collections or highlight

the value of these materials, which in turn presents staffing, system infrastructure, and long-term funding challenges.

The grant project Developing a Framework for Measuring Reuse of Digital Objects, an Institute of Museum and Library Services (IMLS) funded project (LG-73-17-0002-17) by the DLF-AIG, is working to address this through an in-depth needs assessment of the digital library community to determine desired functionality for the construction of a reuse assessment toolkit. The development of well-defined functional requirements and use cases will serve as building blocks that go beyond use and focus on transformation.

The first step of the Developing a Framework for Measuring Reuse of Digital Objects project involved a survey identifying how cultural heritage organizations currently assess digital library reuse, barriers to assessing reuse, and community priorities for potential solutions and next steps.

## 2 SURVEY INFORMATION

The 19 question Qualtrics survey was distributed to 25 different listservs and advertised via the project group's website and the Digital Library Federation's social media channels. The survey remained open from September 13 – October 11, 2017.

Ten questions constitute the primary question set. Questions included those related to institution size, employment, and mission, and collection of both use and reuse data [3].

## 3 RESPONSES

In total, 409 surveys were started. Of those, 107 surveys were eliminated from this data set because the respondent indicated they did not agree to the IRB statement or the respondent did not answer the IRB statement agreement. In both of these scenarios, respondents were not shown any other questions in the survey. The remaining 302 surveys had at least one answer beyond the IRB agreement. This data set (which will be referred to as the total response set) represents 73.8% of all surveys started. 43% (n=130) answered all of the basic questions from the total response set.

Most respondents were from libraries in academic institutions. 30% of respondents reported being from an institution that served under-represented groups.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

JCDL '18, June 3–7, 2018, Fort Worth, TX, USA

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5178-2/18/06.

<https://doi.org/10.1145/3197026.3203871>

## 4 ANALYSIS

### 4.1 Measuring Use

80% of respondents reported collecting use statistics. Google Analytics and platform specific statistics were the most commonly used data-collection tools. The three most common types of use statistics gathered were: number of visitors (12.59%), number of downloads (10.75%), and number of clicks (9.49%). The most commonly reported reasons for collecting use statistics were to "improve digital collection services" (20.19%) and "demonstrate the impact of collections on scholarship and research" (17.13%). The least commonly reported reason for collecting use statistics was to "locate pirated content for intellectual property purposes" (3.32%).

Respondents who did not collect use statistics listed the following reasons: lack of digital collections to measure, lack of staff or staff time, the need for external departments/units to collect the data, lack of priority for usage assessment, lack of tools and systems in place for gathering data, lack of policies regarding data collection, the need for concrete use cases for collecting data, uncertainty about what to do with statistics data after it has been collected, and the new or nascent stages of the respondent's digital collections.

### 4.2 Measuring Reuse

40% of respondents indicated that they collected reuse statistics. The most common methods for collecting reuse data were "social media metrics" (22.32%), "reverse image lookup" (13.30%) and "alert services" (13.73%). The most common kinds of reuse data collected were "digital collections and objects cited in scholarship" (19.32%) and "published or re-posted digital objects in digital media" (18.56%). The least common kinds of reuse data collected were "published data sets used in new research" (6.44%). A lack of accepted methodology or systems for gathering reuse data were the reasons provided most frequently for not gathering reuse data. Similar to the reasons for not collecting use statistics, respondents also indicated that a lack of staff or staff time was a significant barrier. Additional reasons included: institutional lack of priority for use or reuse data, difficulty in measuring reuse, uncertainty as to what reuse data was and why it should be collected, concern about privacy and patron anonymity, lack of definitions for and quantitative methods for calculating reuse, and concern that tracking reuse would be inaccurate without being able to calculate full extent of reuse.

### 4.3 Supporting Reuse Data Collection

In order to implement reuse assessment, respondents most commonly reported the need for the following supports: "documented standards and best practices" (14.86%), "personnel" (12.8%), "online tutorials" (12.06%), "open source tools" (11.78%), and "money" (11.68%). Additionally, respondents commented in a free text question that "research in the validity of such measurements" would be important as well as "easily implemented systems" and "better tracking tools for our implemented persistent IDs (handles, DOI)." Personnel was the most consistently rated high-need support across institutions.

### 4.4 Greatest Barriers to Assessment

Respondents overall expressed concern about being able to adequately measure both use and reuse in the absence of professional standards in the field. Lack of staff, time, and resources were key reasons that respondent's institutions were not pursuing more robust assessment activities. They also discussed the inability to interpret and understand data due to a lack of lack of training, understanding of statistics and data, and reliable measurements. For instance, respondents noted that the system-supplied assessment data generated by organizational content management systems differed from assessment tools like Google Analytics to such a high degree that they did not know which data to rely on. Many respondents identified the difficulty of assessing content across platforms with available tools. Additionally, the theme of not knowing what to do with the information that is gathered was expressed in several different ways. Some respondents noted that they did not know if the information they were gathering was valid or reliably interpreted and others expressed concern about not knowing how to turn that data into actionable information. Respondents also expressed that collected data may not be used or appreciated by upper administration and may even be seen as a waste of time.

## 5 PROGRESS

Data from the survey has been used to inform the continued work of the project. The survey results have thus far helped in developing discussion topics for the focus groups such as refining definitions and examples of use and reuse; determining necessary technology and standards for a reuse toolkit; and examining the cultural and ethical implications of reuse assessment. Data from the survey and focus groups will be used to create use cases which will then be assessed in terms of usefulness via a follow-up survey.

## ACKNOWLEDGMENTS

This work is supported by the Institute of Museum and Library Services (IMLS) (LG-73-17-0002-17).

## REFERENCES

- [1] Krystyna K Matusiak. 2012. Perceptions of usability and usefulness of digital libraries. *International journal of humanities and arts computing* 6, 1-2 (2012), 133–147.
- [2] Santi Thompson, Joyce Chapman, Jody DeRidder, Megan Hurst, Elizabeth Joan Kelly, Martha Kyriallidou, Caroline Muglia, Genya O'Gara, Ayla Stein, Rachel Trent, et al. 2015. Surveying the landscape: Use and usability assessment of digital libraries. *Digital Library Federation Assessment Interest Group, User Studies Working Group* (2015). <https://doi.org/10.17605/OSF.IO/9NBQG>
- [3] Santi Thompson, Elizabeth Joan Kelly, Caroline Muglia, Genya O'Gara, Ayla Stein, and Liz Woolcott. 2017. Developing a Framework for Measuring Reuse of Digital Objects A Survey. *Digital Library Federation Assessment Interest Group* (2017). <https://osf.io/ptvh5/>