Developing best practices and ethical guidelines for assessing reuse of digital content

Caroline Muglia (she/hers), University of Southern California
Joyce Chapman (she/hers), Duke University

Library Assessment Conference
March 17, 2021
I. Overview of grant, project team, and deliverables
II. Ethical guidelines
III. Best practices
IV. Next steps
V. Questions!
Digital Content Reuse Assessment Framework Toolkit (D-CRAFT)
Developed by the DLF AIG Content Reuse Working Group (starting in 2015)
July 1, 2019 - December 31, 2021
Ethical guidelines, recommended practices, and assessment training for GLAMRD* institutions

*Galleries, Libraries, Archives, Museums, Repositories, and Data
Elizabeth Joan Kelly, Loyola University, New Orleans

Ayla Stein Kenfield, University of Illinois at Urbana-Champaign

Kinza Masood, Mountain West Digital Library

Caroline Muglia, University of Southern California

Ali Shiri, University of Alberta

Santi Thompson (PI), University of Houston

Liz Woolcott, Utah State University

Joyce Chapman, Assessment Consultant, Duke University

Derrick Jefferson, Diversity Consultant, American University

Myrna E. Morales, Consultant, Massachusetts Coalition of Domestic Workers; PhD LIS candidate, University of Illinois iSchool

Accessibility Consultant, Future Hire

Instructional Design Consultant, Future Hire
ADVISORY BOARD

- Paige Dansinger, Better World Museum
- LaToya Devezin, Jimmy Carter Presidential Library and Museum
- Genya O’ Gara, Virtual Library of Virginia (VIVA)
- Anna Naruta-Moya, Indigenous Digital Archive
- Kelly Riddle, LA County Library
- Betty Rozum, Utah State University
- Kayla Siddell, Xavier University of Louisiana
- Holly Smith, Spelman College
ETHICAL GUIDELINES
FOR ASSESSING REUSE OF DIGITAL CONTENT
What are the Guidelines and why are they important?

- Set of principles to guide an ethical assessment of reuse
- Use case developed from 2017-2019 grant (Measuring Reuse)
- Reviewed professional GLAMRD codes, policies, guidelines
- Consulted privacy and ethics working groups focused on GLAMRD institutions, leveraged Privacy and DEIA consultants
- Goal: Utility! High level goal, scope, terminology and core values with descriptions, examples, practical applications
ASSESSMENT

- Developed a gauge
- Collected codes via Zotero
- 2 groups to read collected documents against the gauge. Every document had multiple readers.
- Voyant to conduct word/phrase frequency analysis
44 codes reviewed
- 4,283 terms reviewed for frequency
- 30 phrases reviewed for frequency
- Using this information, we began to develop Ethical Guidelines

<table>
<thead>
<tr>
<th>Term</th>
<th>RawFrequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethic/ethics/ethical/ethically</td>
<td>1314</td>
</tr>
<tr>
<td>data('s)</td>
<td>1101</td>
</tr>
<tr>
<td>library/libraries</td>
<td>974</td>
</tr>
<tr>
<td>information</td>
<td>920</td>
</tr>
<tr>
<td>research</td>
<td>657</td>
</tr>
<tr>
<td>privacy</td>
<td>647</td>
</tr>
<tr>
<td>code</td>
<td>593</td>
</tr>
<tr>
<td>professional</td>
<td>519</td>
</tr>
<tr>
<td>use</td>
<td>481</td>
</tr>
<tr>
<td>practice(s)</td>
<td>424</td>
</tr>
<tr>
<td>community(ies)</td>
<td>393</td>
</tr>
<tr>
<td>access</td>
<td>343</td>
</tr>
<tr>
<td>counselors</td>
<td>343</td>
</tr>
</tbody>
</table>
Inclusion, Diversity, Equity, Accessibility and Social Justice (IDEAS)

Privacy

Impartiality

Professional Development/ Training

Transparency

Traditional Knowledge, Cultural Heritage, and Intellectual Property
the assessment of use and reuse of
digital content

Drafted by the Digital Content Reuse Assessment Framework Toolkit (D-CRAFT) Project Team and Consultants.

Finalized on November 30, 2020

Here is a refresher on the Commenting feature on the Google Document platform. Here is a short Qualtrics survey for those who prefer to comment on this platform.

Introduction

The work of Galleries, Libraries, Archives, Museums, Repositories, and Data (GLAMR) institutions in 2020, perhaps more than ever before, must be informed by a critical awareness of identity, socio-economic and racial privilege, access, and privacy, among other important criteria. Ethical considerations should govern exchanges between practitioners, users, and communities that may create or donate digital content. These considerations can signal an investment in meaningful collaborations that empower all stakeholders.

The Digital Content Reuse Assessment Framework Toolkit (D-CRAFT) Project Team ("Project Team") developed the Ethical Considerations and Guidelines for the Assessment of Use and Reuse of Digital Content ("Guidelines"). D-CRAFT is a 2.5-year Institute of Museum and Library Sciences (IMLS) federal grant with goals of developing resources, recommended practices, use cases for sustainably measuring and evaluating the reuse of digital assets held by cultural heritage knowledge organizations.1 D-CRAFT is informed by the Project Team's initial 2017-2018 IMLS grant, "Developing a Framework for Measuring Reuse of Digital Objects," which succeeded in creating a toolkit with resources, best practices, and use cases for study of the reuse of digital assets held by libraries.1

Core Values

The Core Values suggested and presented in this document are informed by the professional codes of ethics put forward by the discrete yet complementary communities of practice that form the GLAMR-sphere: library, archival, and museum workers, information and data curators/maintainers.9

The Core Values govern the ways in which practitioners collect, use, and reuse data, and the privacy and confidentiality of collection users when assessing use and reuse of digital objects.


Responsibility of the Practitioner: Pragmatic approaches for practitioners to realize the guidelines when planning and engaging in use and reuse assessment, including analysis and distribution of results.

Transparency: An ongoing, iterative method to maintain responsible, inclusive, and accessible forms of communication to community stakeholders including creators, individuals and communities who may be depicted in collection materials, patrons, and curators.
Advisory Board Members reviewed in 2020; Team implemented comments

Open commenting period: January 4 - February 12, 2021

Sent to 20 lists/ email of groups or organizations led by or including underrepresented groups

Sent to 35 lists/ email of groups or organizations focused on GLAMR

Google comments: 37 individual commenters

Qualtrics: 8 responses
“There is a massive gap... about GLAMR's responsibility to use/reuse digital content that meet the needs of people with functional impairments or disabilities.”

“The privacy section is focused on individual privacy which makes sense if the scope of this is web analytics. If the scope is also assessment for the ethics of sharing content online, then this section should include cultural and community privacy.”

“This is the biggest weak spot of this document: there are many words and phrases in this document that are relatively new in academic writing, which are poorly defined here. I had several moments of frustration as I read, wondering what certain words and phrases mean to the authors, while knowing these words and phrases may mean something different (often something fraught) to the communities with which I work.”
Asked for and received **thoughtful, insightful, meaningful engagement** with community practitioners. Now we have to do something with it!

**Developed a Summary Document** to capture the initial feedback.

**Identifying experts in areas** of Privacy, Traditional Knowledge, Accessibility, and others for a more substantive contribution to the Guidelines.
BEST PRACTICES
FOR ASSESSING REUSE OF DIGITAL CONTENT
## METHODS INCLUDED

### Data collection methods

<table>
<thead>
<tr>
<th>Alert services</th>
<th>Altmetrics</th>
<th>Citation analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus groups</td>
<td>Interviews</td>
<td>Link analysis</td>
</tr>
<tr>
<td>Point of use surveys</td>
<td>Reverse image lookup</td>
<td>Surveys</td>
</tr>
<tr>
<td>Web analytics</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
INFO ON EACH METHOD

DEFINITION
1-2 sentences explaining what this method is used for and how it applies to use/reuse assessment.

USES
Several paragraphs explaining the insights this method can provide for assessing digital object use or reuse in particular.

TOOLS
Up to five common tools that can be used to collect data or perform this assessment method (linked to Tool pages).

RESPONSIBLE PRACTICE
Explain how the Ethical Guidelines applies to this method and whether there are any special concerns.

STRENGTHS & WEAKNESSES
What is this method particularly good at / bad at telling practitioners about the use/reuse of digital content?

HIGHLIGHTED EXAMPLES
Annotated links to relevant articles, case studies, blog posts, etc., that demonstrate use of this method.
TOOLS FOR EACH METHOD

METHOD
- Reverse Image Lookup

TOOLS
- Bing Image Match
- Google Image Search
- TinEye MatchEngine
- Yandex
EXAMPLES OF TOOLS

**TinEye MatchEngine**
*Data Collection Tool*

**Basic information**
- URL: [https://services.tineye.com/MatchEngine](https://services.tineye.com/MatchEngine)
- Cost: Free/Subscription
- Open Source: No

**Used for the following methods**
- Reverse Image Lookup

**How to use this tool for use/reuse assessment**
TinEye is an image search and recognition company. Its MatchEngine is used for its Reverse Image Search by allowing practitioners to compare images placed on other websites. MatchEngine goes beyond Reverse Image Lookup by enabling practitioners the ability to do multiple or batch queries at one time. As a fee-based service, MatchEngine will generate an index of all images that a practitioner wishes to search, stores this index on its servers, and makes the index available depending on the service package purchased by the practitioner.

**Ethical guidelines**
Practitioners should follow the practices laid out in the "Ethical considerations and guidelines for the assessment of use and reuse of digital content." The Guidelines are meant to inform practitioners in their decision-making, and to model for users what they can expect from those who steward digital collections.

**Strengths**
- Alert Services
  - Unlike some other alert services, like Google Alerts and Google Scholar, Mentions include reach, influencers, sentiment, and more, whereas other alert services may be individually and the practitioner must perform their own aggregation.
  - Very customizable and has dedicated tech support.

**Weaknesses**
- Alert Services
  - Functionality varies depending on which plan the institution has purchased. For example, data export is only available starting at the Company plan.

**SocSciBot**
*Data Collection Tool*

**Basic information**
- URL: [http://socsciobot.wlv.ac.uk/index.html](http://socsciobot.wlv.ac.uk/index.html)
- Cost: Free
- Open Source: No

**Used for the following methods**
- Link Analysis

**How to use this tool for use/reuse assessment**
SocSciBot is a free Windows application that crawls a website specified by a practitioner and extracts hyperlinks from it. The program outputs the results in a variety of ways, such as a list of page and link counts, a list of all external links, and a list of directory, domain, and file interlinking. The program allows a practitioner to review output data in a text file, in an Excel spreadsheet, or in a network visualization. This tool can help practitioners identify and quantify the variety of hyperlinks contained on websites, including links to digital objects.

**Ethical guidelines**
Practitioners should follow the practices laid out in the "Ethical considerations and guidelines for the assessment of use and reuse of digital content." The Guidelines are meant to inform practitioners in their decision-making, and to model for users what they can expect from those who steward digital collections.

**Additional guidelines for responsible practice**
It is the responsibility of the practitioner not to overload the target web servers by crawling them too frequently. SocSciBot asks practitioners not to crawl web sites of organizations that may not be able to afford the additional bandwidth, taking into particular account the
We need the suggestions and opinions of experts like you! Please take time to respond to the call for feedback that will be coming later this spring.
LOOKING AHEAD

TO FUTURE GRANT ACTIVITIES
WHAT'S NEXT?

- Ethical guidelines
- Data collection methods & tools
- Educational tools & training
- Hire Instructional Designer Consultant
- Hire Accessibility Consultant
- Toolkit development on DLF dashboard
- Populate D-CRAFT Toolkit with Project Deliverables
THANKS!

Questions?

Project website: https://reuse.diglib.org/
#digreuse

Caroline: muglia@usc.edu
Joyce: joyce.chapman@duke.edu
Initial definitions devised and debated during the Measuring Reuse IMLS-funded project (LG-73-17-0002-17)

Reuse as…
- Indicator of impact
- An action occurring outside the repository
- Transformative act
Use-Reuse Matrix: Use and Reuse Distinctions

- Access
- Consumption
- Reformatted
- Sharing
- Reproducibility
- Enhancement
- Recontextualization
- Transformation
<table>
<thead>
<tr>
<th></th>
<th>Use</th>
<th>Reuse</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Passive interaction</strong></td>
<td><em>Use and Reuse Instances</em></td>
<td><strong>Active interaction</strong> with a digital object(s) that demonstrates an interest or value to an external user</td>
</tr>
<tr>
<td><strong>Access</strong></td>
<td><strong>To come into contact with a digital object</strong></td>
<td><strong>Reuse</strong></td>
</tr>
<tr>
<td></td>
<td>● Browsing digital repositories for content</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Clicking a link for a digital object</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Downloading digital objects</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Accessing a web archive</td>
<td></td>
</tr>
<tr>
<td><strong>Consumption</strong></td>
<td><strong>To view, read, listen, or expose oneself to the intellectual content of a digital object</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Watching a video online</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Reading an article</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Viewing a photograph</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Listening to a song</td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>Reuse</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td><strong>Passive interaction</strong> with a digital object that indicates <strong>potential interest and/or value</strong> to an external user</td>
<td><strong>Active interaction</strong> with a digital object(s) that demonstrates <strong>an interest or value</strong> to an external user</td>
<td></td>
</tr>
</tbody>
</table>

**Recontextualization**

*To alter the surroundings or space that affect the meaning, purpose, or intent of a digital object.*

- Aggregating metadata in a discovery tool

**Reuse Instances**

- Curating sets of digital material, such as People of Color in Medieval European Art History [https://medievalpoc.tumblr.com/](https://medievalpoc.tumblr.com/)
- Curriculum planning K-12 education, e.g. DocsTeach, LOC Teaching with Primary Sources, etc.
- Creating a Pinterest board of digital objects
- Citing a digital object in a scholarly article and adding interpretation
  … among others