

Ethical Considerations and Guidelines for the Assessment of Use and Reuse of Digital Objects

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

Finalized in June 2023

Introduction

The present-day work of Galleries, Libraries, Archives, Museums, and Repositories (GLAMR) institutions 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 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 (“the Guidelines”). D-CRAFT is a multi-year Institute of Museum and Library Sciences (IMLS) federal grant funded project with goals of developing resources, recommended practices 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 use cases for a toolkit that would include resources and recommended practices for assessing the reuse of digital assets held by galleries, libraries, archives, museums, and repositories.²

The Project Team developed the Guidelines to consider the social and political influences of those working in and impacted by GLAMR environments and the importance of aligning ethics, values, and accountability in the work process. As the world faces the global pandemic precipitated by COVID-19 that is ravaging Black and Brown communities at disproportionate rates and as the United States faces the enduring and virulent white supremacy that led to the murder of George Floyd and many other Black people, the Project Team and practitioners must

¹ IMLS Grant, “D-CRAFT:” <https://www.ims.gov/grants/awarded/lq-36-19-0036-19>

² IMLS Grant, “Measuring Reuse:” <https://www.ims.gov/grants/awarded/lq-73-17-0002-17>; “Measuring Reuse, 2017-2018”: <https://reuse.diglib.org/project-history/>

recognize our collective role in the explicit and implicit structures that continue to oppress, marginalize, and dehumanize historically and newly minoritized communities.³

GLAMR institutions provide no shortage of examples related to the oppression of historically and newly minoritized communities, including centering whiteness in selection, acquisition, organization, and development of physical and digital collections; participation in the development and perpetuation of white supremacist, colonial, and flawed knowledge organization systems; and exclusionary digitization practices resulting in biased decision-making and discriminatory digitization of collections.⁴ For these reasons, the Guidelines are designed for utility beyond the D-CRAFT grant that drove their development.⁵

The Goal of the Ethical Guidelines

The Guidelines are intended for practitioners assessing the use and reuse of digital objects. Within the Guidelines, the term “practitioner” refers to an individual who curates, maintains, or preserves digital objects in a digital environment such as a digital library, digital repository, digital museum collection, digital exhibit, data repository, or digital archive.

The Guidelines are intended to inform practitioners in their decision-making when assessing use and reuse. Integral to creating these guidelines are user privacy considerations and centering the concerns and ideas of historically and newly minoritized communities. The Guidelines focus on *assessment*, rather than how digital content should be ethically used and reused.

By grounding assessment within these Guidelines, practitioners may need to pivot away from assessment strategies that were not centered in or informed by privacy, inclusion, and equity. For novice practitioners especially, this can be difficult to navigate within GLAMR institutions

³ This document uses the term “historically and newly minoritized communities” to represent a broad and diverse range of individuals, identities, and communities. Initially, the Project Team employed the term “Black, Indigenous, and People Of Color (BIPOC)” as representative of these communities. Some have interrogated the term “BIPOC” as not being explicitly inclusive of certain communities or “othering” Black and Indigenous communities rather than centering them. Others embrace this term as encompassing and inclusive. To ensure that the Project Team and the Guidelines are as inclusive as possible of all communities, including working class and poor, LGBTQIA2S+, disabled, neurodivergent, and many more, the Guidelines employ the broader term.

⁴ This broad topic has been meaningfully considered by many scholars and community members: Olson, H. A. (2013). *The power to name: locating the limits of subject representation in libraries*. Springer Science & Business Media; Howard, S. A., & Knowlton, S. A. (2018). Browsing through bias: The Library of Congress classification and subject headings for African American studies and LGBTQIA studies. *Library Trends*, 67(1), 74-88; Tomren, H. (2003). Classification, bias, and American Indian materials. *Unpublished work, San Jose State University, San Jose, California*; Bone, C., & Lougheed, B. (2018). Library of Congress subject headings related to indigenous peoples: changing LCSH for use in a Canadian archival context. *Cataloging & Classification Quarterly*, 56(1), 83-95; Baucom, E. (2018). An exploration into archival descriptions of LGBTQ materials. *The American Archivist*, 81(1), 65-83; Christensen, B. (2008). Minoritization vs. universalization: Lesbianism and male homosexuality in LCSH and LCC. *Knowledge Organization*, 35(4), 229-238.

⁵ The current Project Team (including consultants) represents members from public and private academic institutions in the United States and Canada.

that have built programs and collections on particular assessment approaches. Thus the Guidelines seek to empower practitioners to revisit strategies for current practices to more fully embrace new considerations and advocate for change in whatever form that takes in their institution.

The Guidelines function as direction for digital library practitioners and stewards of content hosted, aggregated, and curated by the GLAMR community. The Guidelines provide a framework to affect positive change in an institution. The examples presented in this document are not comprehensive of the entire GLAMR field.

Scope

The scope of these Guidelines encompasses the act of *assessing* use or reuse of digital objects, which is distinct from using or reusing digital objects. The Project Team defines Use and Reuse as follows:

- Use: Passive interaction with a digital object that indicates potential interest and/or value to an external user
- Reuse: Active interaction with a digital object(s) that demonstrates an interest or value to an external user

These definitions, along with eight “levels of engagement,” form the Use-Reuse matrix, which is a model for defining and determining which interactions with digital objects should be considered use and which should be considered reuse. A complete explanation of the Use-Reuse matrix, including development of the definitions, is out of scope for this document but is available elsewhere.⁶

Terminology

- **Accessibility:** For purposes of this document, the term refers to web accessibility or accessibility in a digital environment.
- **ALA:** American Library Association
- **Cisgender:** people who identify with their assigned-at-birth gender.
- **Cisnet:** A portmanteau of ‘cisgender’ and ‘heterosexual’.
- **Core Values:** Concepts put forward in this document that must be considered when ethically assessing the use and reuse of digital library materials.
- **Digital Objects:** Digitized or born digital materials and associated metadata that are collected, preserved, and disseminated by digital libraries. These objects may be, but are not limited to: cultural heritage artifacts, research outputs, scholarship, and data.

⁶ See Kenfield, A. S., Woolcott, L., Thompson, S., Kelly, E. J., Shiri, A., Muglia, C., Masood, K., Chapman, J., Jefferson, D., & Morales, M. E. (2022). Toward a definition of digital object reuse. *Digital Library Perspectives*, 38(3), 378–394. <https://doi.org/10.1108/DLP-06-2021-0044>

- **DPLA:** Digital Public Library of America
- **GDPR:** [General Data Protection Regulation](#) is a European Union data privacy law.
- **GLAMR:** Galleries, Libraries, Archives, Museums, and Repositories
- **The Guidelines:** This document, Ethical Considerations and Guidelines for the Assessment of Use and Reuse of Digital Content, referred to as “Ethical Assessment Guidelines,” presents a framework for assessing digital objects’ use and reuse based on the Core Values.
- **Heterosexual identity:** “Term used to describe the sexual identity of individuals who are romantically and/or sexually attracted to people assigned the “opposite” sex at birth.”⁷
- **IDEAS:** Inclusion, Diversity, Equity, Accessibility, Social Justice
- **Neurotypical:** “Showing patterns of thought and behavior that are typical of most people.”⁸
- **Practitioner:** An individual who curates, maintains, and/or preserves digital objects in a digital environment such as a digital library, digital repository, digital museum collection, digital exhibit, data repository, or digital archive. While often a formal professional position, enthusiastic citizens, novices, hobbyists, and self-taught individuals are also practitioners, especially for community archives, citizen science, and ecomuseum collections.
- **Privacy:** The concept of “the right to be let alone,” privacy refers to the freedom from intrusion or interference from others. Privacy also refers to the ability to control one’s personal data that is collected, used, shared, and retained by others. A lack of privacy can lead to serious negative consequences for the person whose data is involved, ranging from emotional and reputational harms to physical and discrimination harms.
- **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.

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

⁷ Heterosexual Identity. (n.d.). In *Homosaurus Vocabulary*. Digital Transgender Archive. Retrieved April 20, 2023, from https://homosaurus.org/v3/homokit0000624?from_q=heterosexual

⁸ Neurotypical. (n.d.). In *Oxford Learner’s Dictionaries*. Oxford University Press. Retrieved April 20, 2023, from <https://www.oxfordlearnersdictionaries.com/us/definition/english/neurotypical>

the GLAMR-sphere: library,⁹ archival,¹⁰ and museum¹¹ workers, information and data curators/maintainers.¹²

The Core Values inform 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. Technology evolves quickly, so practitioners should expect frequent changes to how best to provide ongoing care for user privacy and best interests. As advocacy of underrepresented populations is paramount, this includes not only awareness of professional obligations but also the responsibility to push these boundaries when the accepted professional codes of ethics, values or other governing documents do not go far enough.

The Guidelines are informed by the following core values and include commentary on “practical applications” of the following values. Practical applications are intended to expand a practitioner’s awareness around these critical topics; to consider the ethics of their practices. For several core values, no clear practical application or solution exists. In these instances, the section is named “practical considerations,” and explores the opportunities for future work and discussion for ways that the concepts can improve the collective practice of GLAMR practitioners.

1. IDEAS (Inclusion, Diversity, Equity, Accessibility, Social Justice)
2. Privacy
3. Traditional Knowledge, Cultural / Heritage, and Intellectual Property
4. Professional Development and Training
5. Transparency
6. Awareness

I. IDEAS

The purviews of voices from historically Western, White, neurotypical, cishet, male, people without disabilities, and non-working-class backgrounds have long been the norm when establishing recommended practices and standards in the West, which remains largely unchanged in the Information Age. Our goal is for the Ethical Assessment Guidelines to re-center embracing true equity, diversity, and inclusion. The practice of re-centering can support

⁹ See the ALA Code of Ethics <http://www.ala.org/tools/ethics> and the IFLA Code of Ethics for Librarians and other Information Workers <https://www.ifla.org/faife/professional-codes-of-ethics-for-librarians>.

¹⁰ See the SAA Core Values Statement and Code of Ethics <https://www2.archivists.org/statements/saa-core-values-statement-and-code-of-ethics> and the ICA Code of Ethics <https://www.ica.org/en/ica-code-ethics>

¹¹ See the AAM Code of Ethics <https://www.aam-us.org/programs/ethics-standards-and-professional-practices/code-of-ethics-for-museums/> and ICOM Code of Ethics: <https://icom.museum/en/resources/standards-guidelines/code-of-ethics/>

¹² See the ASIS&T Professional Guidelines and the ACM Code of Ethics <https://www.acm.org/code-of-ethics>

best practices that value all people and their contributions, regardless of race, ethnicity, tribal affiliation or recognition, nationality, immigration status, gender identity or expression, sexual orientation or identity, economic background, class, age, ability or disability, or language.

Practical Considerations

Practitioners should be aware of – and practice – ongoing self-reflection of their role in promoting or suppressing diversity and inclusion of GLAMR institutions. The goal for practitioners is to amplify perspectives and voices that have been historically or newly underrepresented when performing assessment. In other words, the goal of this section is for practitioners to incorporate IDEAS into the daily work processes and develop (or refine) their personal awareness of these identity areas.

The term “diversity” is used in this document to describe the range of human differences including, but not limited to: group and social differences related to race, ethnicity, socio-economic status, class, gender identity, sexual orientation, ability or disability, indigeneity, national origin, citizenship or immigration status, language, and cultural, political, religious, and other affiliations. An ongoing examination of human diversity as well as differences in education, personalities, skill sets, experiences, and knowledge bases is critical to ensure proper representation of the full spectrum of humanity.

“Inclusion” is used to describe the active, intentional, and ongoing engagement with diversity (as described above) and ensure respect for those differences. We embrace new perspectives on cultures, beliefs, practices, and contributions and are inclusive of those who are newly or have historically been limited, excluded and minoritized.

Power dynamics contribute to the further marginalization of the communities identified above by preventing them from being fully taken into consideration compared to those in positions of power and leadership. We assert to not just be mindful of the privilege of those who possess it noting its use in situations with both conscious and unconscious bias, but to center those typically on the periphery of leadership positions, to flatten bygone patriarchal structures, and treat everyone involved with equality and respect regardless of origins.

Equally important is the critical examination of the values related to social justice, its theory and its practice which includes, but is not limited to, acknowledging the divisions in our humanity related to racism, violence, literacy, human rights, poverty, hunger, and conservation of the environment.

Major systems of oppression work collectively. Examples of these systems include patriarchy, white supremacy, racism, sexism, ableism, economic hierarchies precipitated by capitalism, and many more. These Guidelines help ensure a just and ethical series of standards that provide a basis of fairness regarding access, justice, equity, diversity, and inclusion for all.

II. Privacy¹³

Privacy is a fundamental human right.¹⁴ In the context of GLAMR, privacy is also seen as a condition necessary for research and creativity.¹⁵ GLAMR practitioners are therefore professionally invested in adhering to the community value of privacy, through such approaches as protecting individual users' personal data, cultivating consent from users, and providing transparency about services and assessments.¹⁶ Against this background, the purpose of D-CRAFT is to facilitate digital object reuse assessment in digital libraries and repositories; however, data collection and assessment should not be undertaken at the expense of users' rights to privacy, nor should collection and assessment come at the expense of informed and explicit user consent. The Guidelines strive to center populations who experience the greatest harms through targeted and systematic surveillance structures and technologies. Sometimes the type of data needed to conduct an assessment will carry considerable privacy risks to users. Different populations will experience different privacy harms as well as the level of harm incurred by individual users.

Because of the potential of connecting an individual to use or reuse, data collection and assessment must consider the possible harms that collecting certain user data can bring. In some cases, these privacy risks can be mitigated through following data privacy and information security best practices, some of which are described below. Practitioners must realize, nevertheless, that there will be data or assessment requirements that carry so much unmitigable risk to user privacy that the collection or assessment cannot be done in an ethical manner. Practitioners are advised to keep this in mind while they assess the potential privacy risks and ethical implications with data collection and assessment during the planning and decision processes.

Data and Privacy

In the context of the Guidelines, data can be personal or non-personal. There are different definitions of personal data, particularly in data privacy regulation. We will refer to the personal data definitions provided by the General Data Protection Regulation (GDPR) and the National Institute of Standards and Technology (NIST). The GDPR's definition of personal data is as follows:

¹³ For additional information on ethics related to privacy and technology, consider Digital Library Federation (DLF) Privacy and Ethics in Technology Working Group: <https://osf.io/bdyvq/>

¹⁴ United Nations. "Universal Declaration of Human Rights." <https://www.un.org/en/about-us/universal-declaration-of-human-rights>

¹⁵ Jones, Kyle ML, et al. "'We're being tracked at all times': Student perspectives of their privacy in relation to learning analytics in higher education." *Journal of the Association for Information Science and Technology* 71.9 (2020): 1044-1059.

¹⁶ Zimmer, Michael. "Librarians' attitudes regarding information and internet privacy." *The Library Quarterly* 84.2 (2014): 123-151.; Newman, Bobbi, and Bonnie Tijerina, editors. *Protecting Patron Privacy: A LITA Guide*. Rowman & Littlefield Publishers, 2017. Lee, Una, and Dann Toliver. *Building Consentful Tech*. 2017.

“Data subjects are identifiable if they can be directly or indirectly identified, especially by reference to an identifier such as a name, an identification number, location data, an online identifier or one of several special characteristics, which expresses the physical, physiological, genetic, mental, commercial, cultural or social identity of these natural persons. In practice, these also include all data which are or can be assigned to a person in any kind of way. For example, the telephone, credit card or personnel number of a person, account data, number plate, appearance, customer number or address are all personal data. Since the definition includes “any information,” one must assume that the term “personal data” should be as broadly interpreted as possible.”¹⁷

The National Institute of Science and Technology (NIST) defines personally identifiable information (PII) as:

“Any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual’s identity, such as name, social security number, date and place of birth, mother’s maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information.”¹⁸

It should also be noted that NIST recommends that organizations consider the sensitivity or organizational/individual risks with data that does not fall under the PII definition.

There are different categories of personal data in both definitions:

- Direct identifiers (name, email address, location)
- Indirect identifiers (age, race/ethnicity, gender identity)
- Behavioral data (data generated by activity, including content use/reuse history and browsing activity)

Numerous research studies have shown that indirect identifiers and behavioral data can be used to identify individuals; therefore, the same care applied to direct identifiers needs to also apply to these types of personal data.¹⁹

De-identification methods such as removing direct identifiers, aggregation, obfuscation, and truncation can mitigate – but not entirely remove – the risk of re-identifying an individual in the

¹⁷ General Data Protection Regulation (GDPR), “Personal Data:” <https://gdpr-info.eu/issues/personal-data/#:~:text=The%20data%20subjects%20are%20identifiable,%2C%20genetic%2C%20mental%2C%20commercial%2C>

¹⁸ E McCallister, T Grance, and K A Scarfone, “Guide to Protecting the Confidentiality of Personally Identifiable Information (PII),” Gaithersburg, MD: National Institute of Standards and Technology, 2010, <https://doi.org/10.6028/NIST.SP.800-122>.

¹⁹ Latanya Sweeney, “Simple Demographics Often Identify People Uniquely,” Data Privacy Working Paper (Carnegie Mellon University, 2000, <https://dataprivacylab.org/projects/identifiability/paper1.pdf>.; Arvind Narayanan and Vitaly Shmatikov, “How To Break Anonymity of the Netflix Prize Dataset,” *ArXiv:Cs/0610105*, November 22, 2007, <http://arxiv.org/abs/cs/0610105>.

dataset. One particular method of de-identification, pseudonymization – the replacement of personal identifiers with an identifier that isn't associated with the identifiable individual – is also not immune from re-identification risks. Anonymization – the breaking of the link between data and the identifiable individual – is highly difficult, if not impossible, to achieve with personal data.²⁰ Practitioners must consider re-identification risks of any de-identification plans before collecting, storing, or making reuse assessment data available. Factors relevant to re-identification include number of individuals in the overall dataset, data elements that overlap with external data sources, small classes of records defined by unique combinations of those data elements, weak pseudonymization methods, and patterns of population overlap between the dataset and external sources.²¹

Data must not be collected for an assessment need that is not clearly defined at the point of collection. More often than not, practitioners will find that the core assessment need can be met with non-personal data or with less privacy-invasive methods. Practitioners should institute policies whereby the institution does not collect personal data for assessment purposes since it is rare that personal data is useful to aggregate reuse assessment. Each institution should also decide, document, and make publicly available information on how the organization protects personal and non-personal data.

Practical Considerations

Assessment Tool Considerations

Many practitioners are faced with dual needs: to perform assessment as well as the need to carefully evaluate the assessment tools being utilized so that they may be used ethically with respect to privacy norms and values of the profession. As mentioned above, there will be times where the risk to user privacy is too great to conduct a particular assessment. A part of this risk calculation is what tools are used to collect and analyze user data for assessment. Some tools available for reuse assessment collect personal data by default. Practitioners should turn off all personal data collection settings; however, there will be times where no option is available to do so. Some examples of personal data collected by reuse assessment tools are as follows.

- 1) Internet Protocol (IP) addresses are tracked by some web analytics tools. In the default setup, Matomo anonymizes IP addresses by masking the final octet of the address, and allows implementers to mask this information further. Prior to the July 01, 2023 release of Google Analytics 4, default Google Analytics settings allow Google to collect the IP address of everyone accessing a website, and an IP anonymization feature had to be set up by the practitioner to prevent Google from tracking the information. In Google Analytics 4, IP addresses are anonymized by masking the final octet of the address, similar to Matomo. Even without IP addresses, analytics tools can still collect

²⁰ Boris Lubarsky. "Re-Identification of 'Anonymized' Data." *Georgetown Law Technology Review*, April 12, 2017. <https://georgetownlawtechreview.org/re-identification-of-anonymized-data/GLTR-04-2017/>.

²¹ <https://doi.org/10.5334/egems.270>

other data that can be used to create a digital fingerprint of the user, such as browser, operating system, and other device information.

2) Survey software such as Qualtrics collects IP addresses and location data for each respondent by default. Additionally, surveys are often distributed to a list of known individuals using their names and email addresses, and by default these data fields are stored alongside individual responses. Practitioners must enable the “Anonymize Responses” setting to remove IP addresses and location data from individual responses in Qualtrics. However, this setting will not de-identify personal data submitted as part of the survey itself, creating a re-identification risk.

In each example the tool has some data collection and de-identification settings, but these settings do not fully de-identify or stop all collection of personal data by the tool. Practitioners must weigh the privacy risks with each assessment tool based on the type of data being collected for assessment and how to mitigate these risks through the combination of privacy features and privacy best practices, such as only collecting the data absolutely necessary for conducting the assessment (or data minimization).

An additional concern are the third parties who host and maintain the tools used in assessment. Practitioners face the dilemma of relying on third-party tools that may not follow these Guidelines. The organizations who create these tools may engage in practices that run contrary to the Guidelines, such as lacking robust data privacy and security practices, or engaging in reselling or sharing data to other third parties without user consent. The GLAMR community should advocate for or fund and develop technologies that gather data while providing privacy protections. While practitioners should negotiate contracts with vendors to strengthen user privacy protections, this is not a guarantee that users will not experience privacy harms from the user of the vendor tool. Practitioners, therefore, should explicitly seek out less privacy-invasive assessment technologies that will not violate user data privacy rights. GLAMR institutions should leverage their collective power in the marketplace to influence vendors to implement changes on the topic of privacy and other areas.

That said, practitioners do not control every tool that can gather data for reuse assessment by others in the organization or external third parties. Examples of methods of reuse assessment data collection in which the practitioners might have no control over include reverse image lookup, link analysis, and citation analysis. Practitioners should explicitly state to users what tools are used, used by who, the reason why these tools are used, and what data collection and tracking is automatically done by each tool as a way to help educate the public on the privacy limitations and risks present in the tools.

Informing Users and User Consent

Practitioners must be intentional in centering minoritized communities in the creation and development of assessment frameworks. In particular practitioners must, to the greatest extent possible, acquire informed consent from users for collection of personal data. The process for

informed consent should be deliberate to ensure that those being assessed are informed about the actual processes happening in data collection and assessment. The process should be developed and reviewed by those who are most at risk for privacy harms, and historically have been most vulnerable to injustices in society. Documentation such as for outreach efforts, the consent process, relationship mapping, and a power analysis²² should be carefully and consistently crafted when planning assessment. These assessment planning materials should be developed and reviewed to support shifting demographics in the community or changes within the digital library collection. The materials should also be reviewed as identities and society change over time. What is created today may be harmful or irrelevant later on.

To the greatest extent possible, practitioners should not collect personal data for reuse assessment. Practitioners should consider a practice of collecting minimum viable data—the least amount of data to make an assessment possible. If personal data is collected, then processes of de-identification, anonymization, or pseudonymization should be considered so as to safeguard the privacy of those people represented in the data. Before engaging in data collection and assessment, practitioners must notify users about and obtain user consent for collection and assessment of their data. In addition to being core principles in many foundational privacy frameworks – such as the Organisation for Economic Co-operation and Development Privacy Principles²³ – data privacy regulations usually require some form of informed consent. For example, GDPR requires that user consent is freely given, specific, informed, and unambiguous. Users can also withdraw this consent at any time.²⁴ Whenever possible, practitioners should only collect user data when the user gives informed consent for collection and assessment.

Once data tracking, use, and privacy policies are created for digital collections assessment, practitioners should share these policies with users. These notifications should be written in plain language, avoiding excessive jargon and legalese. The information should be presented to the user on a public website (in the form of a privacy notice). Practitioners should also notify users at the point before data collection occurs. For example, practitioners may post a notification alerting users at the point of downloading that reuse of digital materials downloaded from the collections may be tracked for assessment purposes. This notification should give the user the choice to not proceed with the download if they do not consent to this tracking. Ideally the user should be able to opt into data tracking instead of being forced to forego a potential use/reuse of the materials to opt out of tracking; however, if this is not possible due to technical limitations practitioners must do due diligence in communicating this limitation and restriction clearly to the user in the privacy notice and before initial use of the digital material.

If a user wants to withdraw their initial opt-in consent of their data being collected and used for reuse assessment, there should be accessible mechanisms in place for them to do so. Any opt-

²² Hagan, Teresa, and David Smail. "Power-Mapping—I. Background and Basic Methodology." *Journal of Community & Applied Social Psychology* 7, no. 4 (1997): 257–67. [https://doi.org/10.1002/\(SICI\)1099-1298\(199709\)7:4<257::AID-CASP428>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1099-1298(199709)7:4<257::AID-CASP428>3.0.CO;2-P).

²³ Organisation for Economic Co-operation and Development Privacy Principles: <http://oecdprivacy.org/>

²⁴ GDPR consent requirements: <https://gdpr.eu/gdpr-consent-requirements/>

out should also trigger deletion of data collected about the user in the past. Under the GDPR, this is known as the right to erasure or the right to be forgotten.²⁵ If an option for aggregated tracking is available, this option should be communicated to the user at the point of data collection. If this option is unavailable, that must also be communicated.

In some cases, gaining informed consent might not be possible; for example, when the identity of a user(s) is unknown, investigating their identity would itself present a violation of that user's privacy. In these cases, we recommend that the practitioner provide a plain-language notification to participants about the assessment benefits and the privacy risks. Assessment tools should always respect the Global Privacy Control (GPC: successor to the Do Not Track flag) in browsers as an explicit opt-out of data collection and tracking by the user. Consent to personal data collection for the purpose of reuse assessment should not be a barrier for use of digital materials, nor should the consent process force users to relinquish power or agency in exchange for the use or reuse of digital materials.

Practitioners should not monetize or resell data collected for reuse assessment, nor share data with third parties that will monetize or resell it. To the greatest extent possible, practitioners should negotiate for removal of a third party's right to monetize or resell user data in third party agreements. Engagement in reselling data exploits minoritized communities while uplifting for-profit interests. The potential for targeted manipulation of truth, whether through misinformation (unintentional sharing of false information), disinformation (intentional sharing of false information), or purposeful silence further disempowers minoritized communities. These symptoms of privacy violations weaken relationships and trust and, if not addressed, can lead to a fracturing of the very communities GLAMRs attempt to strengthen.

D-CRAFT encourages practitioners to continue to engage critically and as a community on this complex topic.²⁶

III. Traditional Knowledge, Cultural Heritage, and Intellectual Property

These Guidelines are intended to facilitate assessment of the reuse of digital objects, including instances of reuse of digital cultural heritage, traditional knowledge, cultural information and data. The way in which we have conceived of digital object reuse is significantly influenced by Western conceptualizations of intellectual property, copyright, and fair use. However, there is a movement to provide effective protection for "traditional knowledge, genetic resources, and traditional cultural expressions (folklore),"²⁷ such as that which is enjoyed by works created within the colonial ideology of intellectual property.

²⁵ GDPR Article 17, "Right to Erasure": <https://gdpr-info.eu/art-17-gdpr/>

²⁶ For additional information on ethics related to privacy and technology, consider Digital Library Federation (DLF) Privacy and Ethics in Technology Working Group: <https://osf.io/bdyvq/>

²⁷ *Traditional Knowledge and Intellectual Property – Background Brief*. (n.d.). World Intellectual Property Organization. Retrieved April 21, 2023, from <https://www.wipo.int/publications/en/details.jsp?id=3858>

Practical Considerations

For the most part, these Guidelines address digital object use and reuse *assessment*. However, ethical guidelines around traditional knowledge, cultural heritage, and intellectual property are tightly intertwined with the acquisition, stewardship, ownership, and cultural origins of these materials, which can impact digitization and assessment. This section therefore includes a discussion of ethics around acquisition and stewardship of materials.

A component of this process must include verification that the acquisition of the materials was conducted in an appropriate manner following Indigenous protocols and confirming these materials were obtained with permission of Indigenous nations. Historically, materials have been acquired without the permission or knowledge of the Indigenous communities to which they belong. In other cases, for example in the case of photographer Edward Curtis,²⁸ Indigenous people were displayed to create the greatest impact regardless of whether or not the image was culturally correct. Most of Curtis' photographs were designed to meet with the "Vanishing Race" theme, which removed any link to modernity and encouraged the inclusion of attire that was not always authentic.

When working with materials from Indigenous and minoritized communities, including in a reuse assessment, additional research into their values, beliefs, and concerns is necessary to establish a trustworthy exchange between the community and the repository presenting, or planning to present, those materials digitally. An assessment may uncover breaches in Indigenous protocols as well as aspects of privacy and equity that must be addressed. Indeed, the very assessment of traditional knowledge and "ways of knowing" is defined differently by each Indigenous community. In order to create and foster a relationship built on trust, these issues must be carefully considered by practitioners.

The idea of assessing digital object use and reuse, beyond respect for and adherence to the specific protocols of the creating community, is a Western concept. Ideally, establishing that materials have been acquired ethically and in accordance with the community's protocols, should be undertaken prior to digitization. However, if situations arise where, during the reuse assessment of digital materials, it is discovered the Indigenous community did not participate, a move to create the appropriate relationship must be undertaken. The decision to continue allowing for the use or reuse of pertinent materials would then lie with the affected Indigenous community.

In an effort to protect Indigenous rights and interests, the International Indigenous Data Sovereignty Interest Group within the Research Data Alliance created the Collective Benefit, Authority to Control, Responsibility, and Ethics (CARE) principles for Indigenous Data governance. These principles drew from the Canadian Ownership, Control, Access and Possession (OCAP) principles which ensure participation and respect of the Indigenous community involved. These principles safeguard the Indigenous communities' stewardship over their cultural materials and belongings. OCAP acknowledges Indigenous ownership, control, access, and possession of all cultural materials and asserts Indigenous control over data collection processes, possession of materials, and control over how this information can be used.

²⁸ Mathilde Arrivé, « Beyond True and False? », *Études photographiques* [En ligne], 29 | 2012, mis en ligne le 18 juin 2014, consulté le 02 février 2023.

<http://journals.openedition.org/etudesphotographiques/3474>

The primary goals of the CARE principles are to (1) foster Indigenous self-determination by enhancing Indigenous use of data for Indigenous pursuits and (2) honor the “Findable, Accessible, Interoperable, and Reusable (FAIR) Guiding Principles for scientific data management and stewardship”²⁹ while ensuring data sharing on Indigenous terms.

Evaluation of materials to be shared digitally must be conducted in partnership with the Indigenous Nation to whom these materials belong. The overarching adage to this practice is “Nothing about us without us.” The Western concept of copyright cannot be applied to materials which by their nature are stewarded by the nations or families within these nations. Individual ownership, especially of those materials considered sacred or which hold deep cultural meaning, is not a concept practiced by most Indigenous nations. Particular imagery, stories, dance and regalia may be held for future generations by families or members within the nations, but are not considered owned by any one person within the community.

Anderson notes, “Indigenous people must be centrally involved in developing appropriate frameworks for access and use of their knowledge and knowledge practices.”³⁰ She articulates how “Indigenous people are asking for their cultural systems and ways of governing knowledge access and use to be recognized as legitimate, and to be respected as custodians/owners/nurturers of knowledge that is valuable within and beyond Indigenous contexts.”

In June 2019, the Canadian government released the Statutory Review of the Copyright Act.³¹ This report was a result of the Standing Committee on Industry, Science and Technology, which reviewed Canadian copyright legislation. As a result of the Indigenous presentations made to the Standing Committee, recommendations were made that included “the recognition of an effective protection of traditional arts and cultural expressions in Canadian law, within and beyond copyright legislation.” Further, the recommendations went on to include “the participation of Indigenous groups in the development of national and international intellectual property law” and “granting Indigenous peoples the authority to manage traditional arts and cultural expressions, notably through the insertion of a non-derogation clause in the Copyright Act.” These recommendations were based on testimony from Indigenous communities. The report recognized, “Contrary to classic conceptions of copyright ownership, which grants individual ownership based on the idea that works originate from one or a few individual authors, for Indigenous witnesses, traditional arts and cultural expressions have communal ownership.”

This communal ownership approach should be considered as part of responsible digital content stewardship. Ensuring Indigenous communities are consulted and involved in the digitization, sharing, and reuse assessment of their cultural materials is paramount to authentic representation and respect for traditional knowledge and Indigenous stewardship. The sharing of high-resolution surrogates of sacred objects that were obtained either illegally or under an unauthenticated chain of provenance is one potential ethical consideration in any assessment that practitioners may encounter. Since the stewarding repository has no control over how the materials were obtained, whether they have been obtained ethically, and what the patron will do

²⁹ Wilkinson et al. 2016. <https://doi.org/10.1038/sdata.2016.18>

³⁰ Anderson, J. (2010). *Indigenous/traditional knowledge & intellectual property*. Duke University School of Law Center for the Study of the Public Domain. https://web.law.duke.edu/cspd/pdf/ip_indigenous-traditionalknowledge.pdf

³¹ Statutory Review of the Copyright Act. 2019. <https://www.ourcommons.ca/DocumentViewer/en/42-1/INDU/report-16/>

with the content afterward, there is the potential that sharing these objects violates sacred protocols of a particular Indigenous community.

Thus, it is the responsibility of the practitioner not only to evaluate intellectual property status, but also to build relationships between the materials and the users. The first step is to determine the provenance. Practitioners should consider some of the following questions: how did the material come to be included in the repository? How are they categorized? Which resources can be used to authenticate its source and categorization?

By engaging with the Indigenous communities who are stewards of the cultural materials and by honoring a principle which asserts their sovereignty over the materials, a practitioner asked to provide digital files to a patron will not unknowingly violate Indigenous intellectual property or copyright laws.

IV. Professional Development and Training

Professional development and training is necessary to gain new skills and hone existing ones, invite new voices that may positively influence a practitioner's perspective, and surface assumptions that may undergird a practitioner's daily operations and work. This is especially important in the context of assessment, as the results of assessment are often used as justification for decision making, including data-driven policy development, digitization priorities, and funding needs. Professional development and training in assessing reuse of digital objects, and doing so ethically, is a new area of research and practice. Thus practitioners must engage both in their professional role as a steward of digital cultural heritage artifacts, as well as with the ethical considerations of assessment of reuse of those artifacts.

Practical Considerations

Practitioners should not attempt to assess reuse without knowledge of the selected methodology, framework, and/or assessment tool(s). Assessment training should follow along the lines of practitioner self-reflection and critical awareness. For instance, training could integrate anti-bias modules to support a practitioner's self-awareness, and other power dynamics – both explicit or inherent. Any policy changes or recommendations based on assessment activities should acknowledge and document the practitioners' level of expertise and familiarity with the selected assessment methodologies and tools. Similarly, employing algorithms and/or artificial intelligence (AI) in assessment should be done with extreme care. Many of these tools are proprietary and their code is not made available to anyone outside of the company that develops them. This is of concern because research has shown that computer algorithms³² and AI often reflect the biases of their human programmers and lack diversity in training datasets,³³ among other issues. AIs and computer algorithms can be helpful tools for

³² For more information see Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

https://openlibrary.org/works/OL19734838W/Algorithms_of_oppression

³³ For more information see Buolamwini, J. & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness,*

practitioners, but must not be considered objective, nor used as the sole method of assessment. Further, practitioners should make available assessment findings that are cited as supporting evidence for decision-making as well as the data, methodologies, tools, and analyses without violating privacy of the involved parties.

V. Transparency

Transparency begins with responsible, inclusive, and accessible forms of communication to community stakeholders including creators, individuals and communities who may be depicted in collection materials, users and patrons, and curators. The Core Values documented in the Guidelines are included in the topic, concept, and application of transparency throughout the digital life cycle. As such, this section stands alone for practitioners specifically interested in the role of transparency in their work, and also informs the other Core Values. To maintain transparency is to communicate the process and policies by which digital library practitioners manage and handle information. Indeed, those processes and policies should consider and uphold principles outlined earlier in the Guidelines related to Privacy and IDEAS.

Practical Application

To ensure maintainable transparency, the Guidelines strongly recommend that practitioners document their process and actions when assessing reuse of digital assets. Documentation such as policies, guidelines, or best practices should be made publicly available and drafted in plainspoken language for the general public with a minimum of jargon. This may also include having documentation translated into other languages. Public availability refers to anyone rather than just the stakeholders immediately invested in the materials or project. Ideologically, offering information publicly aligns with the Guidelines. Technically, some digital library practitioners may find it difficult to identify or locate all affected populations. One possibility for addressing this is to provide links to these policies on the same webpage(s) as the digital objects themselves.

Availability of documentation may include translations in multiple languages, in digitally accessible formats, using non-proprietary software to draft and store documentation, and/or providing report summaries or overviews of lengthier documentation protocols to ensure all community members have entry points to interact with the information. Many institutions and organizations share accessibility standards, which may be followed.³⁴ Documentation, such as a change log,³⁵ should also be maintained for modifications and updates to a process or protocol

Accountability and Transparency, in *Proceedings of Machine Learning Research* 81:77-91 Available from <https://proceedings.mlr.press/v81/buolamwini18a.html>.

³⁴ Examples include: Digital Library Federation, “Digital Accessibility Group:” https://wiki.diglib.org/Digital_Accessibility_Group; Society of American Archivists, “Guidelines for Accessible Archives for People with Disabilities:” <https://www2.archivists.org/groups/reference-access-and-outreach-section/guidelines-for-accessible-archives-for-people-with-disabilities>

³⁵ A change log (or changelog, both spellings are used) is a record of edits, including additions, revisions, and deletions made to a digital object (e.g. website, electronic document, software code). It can be thought of as a type of provenance metadata. [Keep a Changelog](#) defines a ‘changelog’ as “a file which contains a curated, chronologically ordered list of notable changes for each version of a project.”

that may impact the information. Thus it is advised to return to existing documentation at intervals most effective for the institution or group and update processes such that they reflect the most current operations, standards, protocols, and guidelines.

If one's workplace does not retain such communication, those forms of communication are not regularly reviewed and revised, and/or if those practices are not upheld, consider this a first step towards transparency. Without a clear understanding of approaches to managing sensitive information or clearly conveying assessment procedures, the concept of transparency will become more difficult to achieve and sustain as a practitioner or as an institution. The governing documents should also reflect transparency. Further, practitioners should make available assessment findings that are cited as supporting evidence for decision-making as well as the data, methodologies, and analyses without violating privacy of the involved parties.

VI. Awareness

It is the position of the Guidelines that humans, or machines and tools created by humans such as computer algorithms and AI, cannot achieve true neutrality, impartiality, or objectivity. All actors are influenced by their environments, circumstances, and lived experiences, which shape every aspect of our lives, the interpersonal and material, including the work of digital library practitioners. When assessing reuse, awareness of one's own intersectional identities and frames of reference should be acknowledged explicitly. The same should be true for any tools or methodologies used to conduct reuse assessment, which may not be possible when using proprietary software, algorithms, or AI.

Rather than striving for "neutrality," the goal should be transparency and equity as defined in this document.

Practical Application

Practitioners should recognize that they are not, cannot be, and should not be neutral or impartial when conducting assessment. Rather, when assessing reuse of digital objects, practitioners have the responsibility to listen to, advocate for, and protect the personal data of any users – particularly members of historically and newly minoritized communities and individuals affected by assessing the reuse data. While this document acknowledges that individuals are not and should not be considered "impartial" or "neutral," it is the responsibility of the practitioner to be cognizant of their own identities. Practitioners should consider the limitations inherent to assessment technologies and account for the impact of biases in these tools before making decisions or recommendations based on reuse assessment.

[CASRAI's Research Data Management glossary](#) provides a more general definition: "Tracks the progress of each change from submission through review, approval, implementation and closure. The log can be managed manually by using a document or spreadsheet, or it can be managed automatically with a software or Web-based tool."

The practitioner should be cognizant of their identities when assessing reuse as well as any policies, guidelines, protocols created as a result of those assessments. This can be expressed transparently through documentation, which may include adding biographical statements.³⁶

Further Reading

- Caswell, M., & Cifor, M. (2016). From Human Rights to Feminist Ethics: Radical Empathy in the Archives. *Archivaria* 81, 23-43. <https://www.muse.jhu.edu/article/687705>.
- Digital Library Federation (DLF) Privacy and Ethics in Technology Working Group: <https://osf.io/bdyvq/>
- Olson, H. A. (2013). *The power to name: locating the limits of subject representation in libraries*. Springer Science & Business Media.
- Howard, S. A., & Knowlton, S. A. (2018). Browsing through bias: the Library of Congress classification and subject headings for African American studies and LGBTQIA studies. *Library Trends*, 67(1), 74-88.
- Tomren, H. (2003). Classification, bias, and American Indian materials. *Unpublished work, San Jose State University, San Jose, California*.
- Bone, C., & Lougheed, B. (2018). Library of Congress subject headings related to indigenous peoples: changing LCSH for use in a Canadian archival context. *Cataloging & Classification Quarterly*, 56(1), 83-95.
- Baucom, E. (2018). An exploration into archival descriptions of LGBTQ materials. *The American Archivist*, 81(1), 65-83.
- Christensen, B. (2008). Minoritization vs. universalization: Lesbianism and male homosexuality in LCSH and LCC. *Knowledge Organization*, 35(4), 229-238.
- World Intellectual Property Organization. "Traditional Knowledge and Intellectual Property – Background Brief." Accessed November 8, 2020. <https://www.wipo.int/publications/en/details.jsp?id=3858&plang=EN>.
- Dressler, V. (2018). *Framing Privacy in Digital Collections with Ethical Decision Making* (San Raphael, CA: Morgan and Claypool). <https://doi.org/10.2200/S00863ED1V01Y201807ICR064>
- Outcomes from National Web Privacy Forum: <https://www.lib.montana.edu/privacy-forum/>
- Hirsch, Dennis D. and Bartley, Tim and Chandrasekaran, Aravind and Parthasarathy, Srinivasan and Turner, Piers and Norris, Davon and Lamont, Keir and Drummond, Christina, Corporate Data Ethics: Data Governance Transformations for the Age of Advanced Analytics and AI (September 10, 2019). Ohio State Public Law Working Paper No. 522. Available at SSRN: <https://ssrn.com/abstract=3478826>

³⁶ An example of inclusion of biographical statements and other forms of transparency of a process: Light, Michelle, and Tom Hyry. "Colophons and Annotations: New Directions for the Finding Aid." *The American Archivist* 65, no. 2 (September 2002): 216–30. <https://doi.org/10.17723/aarc.65.2.13h27j5x8716586q>.

Acknowledgements

We would like to thank our paid [experts](#) who made significant written contributions to the following sections:

Monique Manatch - Traditional Knowledge, Cultural Heritage, and Intellectual Property

Becky Yoose - User Privacy

Scott Young - Data Privacy