



Journal of Copyright in Education and Librarianship

ISSN 2473-8336 | jcel-pub.org

Volume 5, Issue 1 (2021)

Formulating a Scalable Approach to Patron-Requested Digitization in Archives

Kevin S. Hawkins and Julie Judkins

Hawkins, K. S. & Judkins, J. (2021). Formulating a Scalable Approach to Patron-Requested Digitization in Archives. *Journal of Copyright in Education and Librarianship*, 5(1), 1-11. <https://doi.org/10.17161/jcel.v5i1.14652>



© 2021 Kevin S. Hawkins and Julie Judkins. This open access article is distributed under a [Creative Commons Attribution- 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Formulating a Scalable Approach to Patron-Requested Digitization in Archives

Kevin S. Hawkins, University Libraries, University of North Texas

Julie Judkins, University Libraries, University of North Texas

Author Note: Correspondence regarding this article should be directed to Kevin S. Hawkins, Assistant Dean for Scholarly Communication, University Libraries, University of North Texas, kevin.s.hawkins@ultraslavonic.info.

Abstract

The novel coronavirus of 2019 (COVID-19) crisis has forced archives to rethink their modes of providing access to physical collections. Whereas difficult copyright questions raised by reproducing items could previously be skirted by requiring researchers to work with materials in person, the long-term closure of reading rooms and decrease in long-distance travel mean that archives need a workflow for handling user digitization requests that is scalable and requires consulting only easily identifiable information and, assuming full reproduction is off the table, reproducing items in a collection under 17 U.S.C. § 108 or through a strategy of rapid risk assessment. There is a challenge in creating a policy that will work across different formats and genres of archival materials, so this article offers some suggestions for how to think about these parameters according to US copyright law and calls for a committee of experts to work out a model policy that could serve remote users of archival collections even after the COVID-19 crisis has passed.



The Problem

Archives, which primarily hold unpublished material in their physical collections, present special problems for those who wish to reproduce items from these collections. The donor of a collection may have placed restrictions on access to or reproduction of items from the collection, or there may be privacy or ethical issues that would lead the archive staff to place such limitations on use. Such needs to impose restrictions require careful record-keeping and policies to ensure appropriate use of collections.

Those wishing to reproduce items from a collection face an additional problem: ownership of copyright. The authorship of unpublished material is often unclear, and even when the authorship is known it can be quite difficult to locate the author, confirm the author's date of death (required for determining the term of copyright), and figure out who currently owns the copyright. Archives typically put responsibility for determining copyright ownership and for operating duplication equipment on users, relying on the safe harbor provided in 17 U.S.C. § 108(f)(1) to avoid liability for any infringement by users.

Yet during COVID-19, most reading rooms closed or had limited access, meaning users were not able to make their own reproductions. User requests to have items digitized and shared without visiting the reading room became more common during the coronavirus pandemic at our institution (10 requests in 2019 but 43 in 2020), and we expect the same is true for other archives. If archive staff wish to continue providing access to collections during this period, they need to develop procedures for determining what they will digitize for patrons. Given the difficulty of confirming copyright status, archive staff often base such procedures on risk assessment.

This article offers a model for addressing patron digitization requests based on a risk assessment according to US copyright law. For cases where there are no restrictions on access to a collection (either imposed by a donor or relating to privacy or ethical issues), we propose a scalable workflow to perform a check on the copyright status of requested items and, in cases where there is not an unambiguous right for the archive to reproduce the work in full, digitization of portions of the work using certain provisions of 17 U.S.C. § 108 or through a rapid risk assessment. The main difficulty here, naturally, is in creating a method for rapid risk assessment that will work across different types of materials.

Literature Review

The University of California Berkeley Library (2020) created and shared workflows discussing four key areas of law and policy relating to digitizing archival collections: donor restrictions, privacy issues related to people portrayed or discussed in a collection, ethical concerns about making certain information available, and copyright. Other workflows focus only on risk assessment for copyright (UCLA Library Special Collections, n.d.; Fialkov, O’Keeffe, & Kaufman, 2017). However, the use case for these workflows is large-scale digitization, not patron-requested digitization by archive staff, so for our purposes they should not be adopted as is.

The dominance of unpublished material in archival collections means that the copyright status of large portions of an archival collection is likely to be uncertain. This uncertainty leads many archivists to avoid digitization when there is any chance of copyright infringement (Smith, 2012; Dryden, 2014), but according to Smith, “copyright should be treated in the same way as other risks of legal liability, as a subject of risk management,” with rational strategies for evaluating the risk of digitizing broad categories of items as part of a large-scale digitization project (p. 17). Smith described (2012; pp. 19–20) four strategies for managing the risk of engaging in a large-scale digitization project of an archival collection:

1. Assessing the likelihood that certain groups of materials are in the public domain.
2. Asking permission from those who would most likely object to online display.
3. Creating a take-down policy by which a member of the public could lodge an objection to making a particular work available online.
4. Considering a fair-use argument that could be used if necessary.

Though these strategies are sensible as approaches to digitizing a whole collection to make it available online to the public, in cases where a researcher asks an archive’s staff to reproduce material from its collection directly for the user, all four strategies may not be appropriate: a take-down policy would not apply, and seeking permission from creators, current copyright owners, or subjects of works who might object to display of the work online could introduce unpredictable delays in responding to the patron request.

What alternative strategies might exist? 17 U.S.C. § 108 specifies “conditions under which specific copying and distribution activities may take place [by library or archive staff] without infringing copyright” (Buttler, 2012, p. 290). Subsections 108(d) and 108(e) specifically address patron-requested reproduction of works held in archival collections by archival staff. Specifically, subsection 108(d) allows reproduction for a patron of a single “contribution to a copyrighted collection” or “a small part of any other copyrighted work,” and subsection 108(e) allows reproduction of an “entire work, or . . . a substantial part of it,” if the archive staff determines that a copy “cannot be obtained at a fair price”—in both cases where archive staff have “no notice that the copy or phonorecord would be used for any purpose other than private study, scholarship, or research.” However, both of those subsections are limited by subsection 108(i), which excludes images, motion pictures, audiovisual works “other than an audiovisual work dealing with the news,” and a few other categories less frequently held in archives from this protection of liability for archives. It also excludes from protection of liability “illustrations, diagrams, or similar adjuncts to [pictorial or graphic] works” reproduced according to the provisions of subsections 108(d) and 108(e). Furthermore, according to subsection 108(a), copies must be made “without any purpose of direct or indirect commercial advantage” and include a notice of copyright, and the archive must be open to the public or to researchers not affiliated with the institution of which it is a part.

A second alternative strategy is *controlled digital lending* (CDL), in which digital surrogates of archival material are shared with a single researcher at a time (Bailey et al., 2018). By maintaining an “owned-to-loaned” ratio, the CDL model can be thought of as allowing archival staff to avoid item-level risk assessment by loaning a work digitally in place of the print version and in a similar manner (to one user at a time). However, significant work is required to implement the digital rights management technology required for CDL (Cheney, 2020). A simpler version of single-user access, a *virtual reading room*, has also been used with special collections material. This model can involve obscuring content from web crawlers to limit discoverability (Light, 2014), thereby strengthening a fair-use argument. The increasing ubiquity of video-sharing technology allows for an alternative approach to sharing material with users without prior digitization: having archive staff livestream from an overhead camera to simulate the experience of viewing material in a reading room.

Fair use has also been offered as a general principle by which archival collections can be digitized—through the transformative use of aggregating unique collections “with commentary, criticism, and other curation” according to a code of best practices from the Association of Research Libraries, Center for Social Media, and Program on Information Justice and Intellectual Property (2012, p. 19). However, as with the workflows provided by UC Berkeley and UCLA, the use case presented in this code of best practices is mass digitization, not one-off patron-requested digitization. While Smith (2012) offered the fair-use doctrine as a last resort in minimizing the risk of large-scale digitization of archival collections, fair use has also been used as a primary method of managing risk, generally in combination with an assessment of the genre of content (Dickson, 2010; Fortney, 2013; Evans & Hernández Durán, 2018). Similarly, the “Public Statement by Library Copyright Specialists: Fair Use & Emergency Remote Teaching & Research” (2020) argues that the public health emergency caused by COVID-19 weighs heavily in favor of fair use to allow for continuity of teaching and research at US colleges and universities. However, the statement also recommends mitigating the risk of infringement through actions such as providing access to content only to students enrolled in a particular course and for the period of time needed. Neither of these applies to a nonstudent researcher who will certainly want to keep a personal copy of the requested material indefinitely.

Despite the many available workflows, statutes, and models archivists and librarians have used to make physical collections available in digital form, we believe that not one is entirely appropriate for a situation in which legal risk must be assessed rapidly for heterogeneous archival materials in archival collections.

Proposed Workflow for Assessing the Risk of Patron Digitization Requests

Below we propose a sketch of a workflow for rapid assessment of the risk of digitizing unpublished items from an archival collection in response to a patron request.



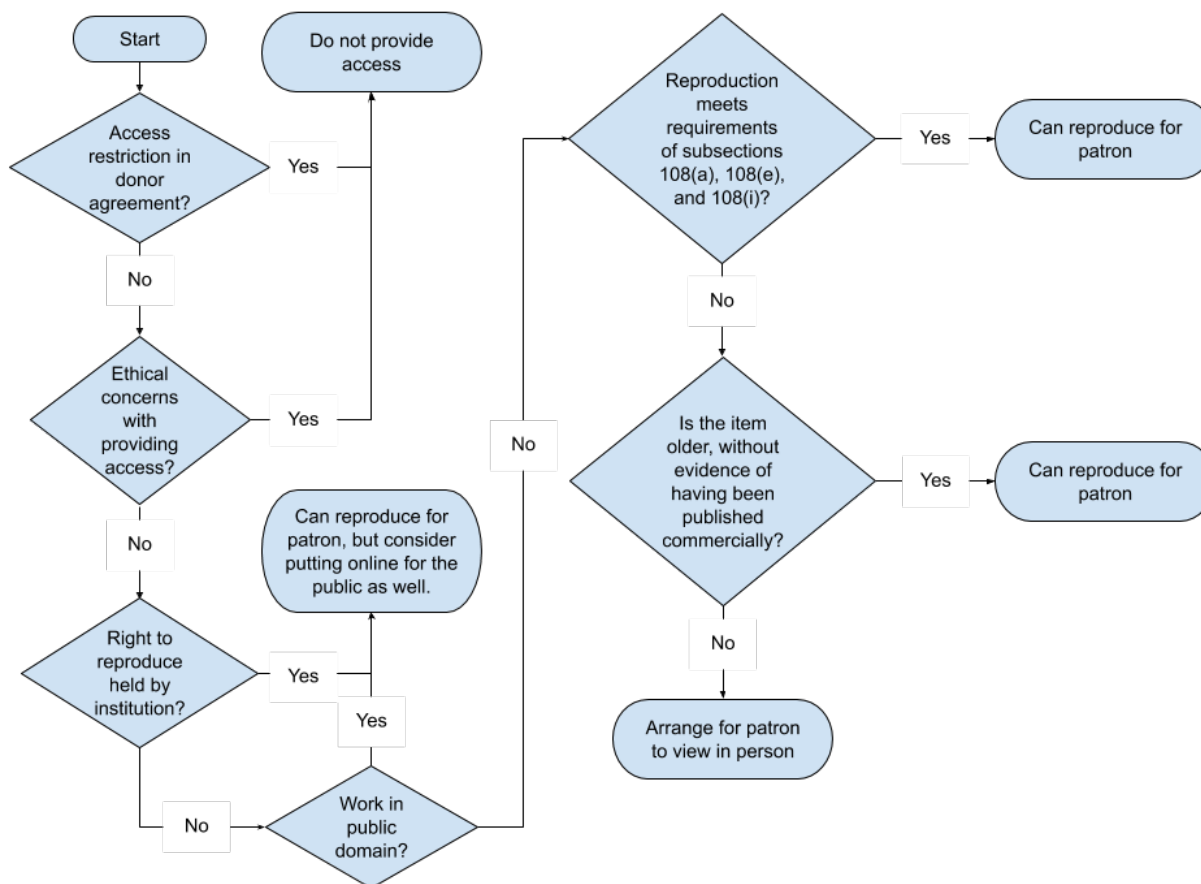


Figure 1. Simplified diagram of workflow for assessing the risk of patron digitization requests.

If a patron requests digitization of one or more items from a collection, the first step is to check the donor agreement (assuming one exists) for any restrictions on access. If the archive has any donor agreements that include a grant of rights to the institution, the donor agreement for the requested item should be checked for this as well since it might allow for digitization for the patron without any further investigation. Assuming the donor agreement allows digitization, and assuming sharing the content does not raise any ethical or privacy concerns as described by the University of California Berkeley Library (2020), archive staff could provide a digital copy of the item to the patron directly through any appropriate means, such as a file-transfer service; if the archive has the staff and technical capacity, the item could also be added to a public portal for sharing archival collections with the world.

Next, archive staff should consult *Copyright Term and the Public Domain in the United States* (Hirtle, 2020) to see if the work(s) contained in the requested item

appear to be in the public domain based on easily identifiable information, such as a date of publication on the item and a date of death of the creator(s) as recorded in the finding aid or donor records. If it is clear that the item is in the public domain, a digital copy of the item could be provided as described above.

Otherwise, the work(s) contained in the requested item is more likely to be protected by copyright, so if proceeding with digitization as described below, the copy of the item should be shared directly with the patron rather than added to a public portal for sharing archival collections.

Assuming the archive meets the requirements of subsection 108(a) of the Copyright Act (1976), including not reproducing or distributing digitally for “any purpose of direct or indirect commercial advantage,”¹ the next step would be for archive staff to verify whether digitization could proceed under the requirements section 108.² While subsection 108(d) is likely unworkable due to the difficulty of mapping the limitations on amount to archival material, subsection 108(e) provides a clear pathway in cases for digitizing any work not excluded in subsection 108(i) where the archive staff determine “that a copy or phonorecord of the copyrighted work cannot be obtained at a fair price” (17 U.S.C. § 108[e]).

If these provisions in section 108 do not apply, then possible alternatives would include using CDL or a virtual reading room. However, CDL will be unsatisfactory to patrons who will in most cases want to keep a copy of the material consulted rather than only view it online during the loan period. And virtual reading rooms, generally used for content digitized in advance whose availability is not meant to be widely advertised, do not seem appropriate for providing access to specific content by request.

Instead, archive staff should undertake a rapid assessment of the risk of digitizing the requested item and providing the patron a copy of the item. We believe that archive staff are not in a position to carry out a rapid fair-use analysis for patron-requested digitization of an item from a heterogeneous collection, especially given that archive staff are likely to have an incomplete understanding of

¹ The American Association of Law Libraries, among others, argues that a “library [or archive] may charge a fee for making the copy as long as the charge does not exceed reasonable cost recovery” (American Association of Law Libraries, 2014).

² The authors are grateful to Kyle K. Courtney, who urged us to consider this to be a viable part of the workflow despite our initial skepticism.



the patron's intended use for the requested item. Instead, factors that could help mitigate legal risk could include:

1. An older work is less likely to have a rightsholder who plans to commercially exploit the work or who would come forward to object to reproduction for a patron request.
2. If there is no evidence on the item itself or in the finding aid that the work embodied by the item was later published commercially, it is unlikely that copyright was previously transferred by the creator to another party, or that the creator granted an exclusive license to another party, either of which would raise the risk that the current rightsholder would object to the digitization.

Thus, we believe that if the requested item is sufficiently³ old (as determined by examining the item or by skimming the contextual information provided in the finding aid), with no evidence that the work was later published commercially, the risk of digitization and providing the patron a copy of the item is low.

In addition to notice of copyright required by 17 U.S.C. § 108(a)(3) for a protection of liability for copies produced by archive staff, it would be prudent for the digitized copy provided to the patron to state that the copy is meant for use only for personal research.

Further Research

We recommend that the Society of American Archivists convene a committee of experts, including those from the Association of Research Libraries and the Center for Media and Social Impact, to issue a model procedure on scalable patron-driven digitization. Most crucially, this model procedure would need to suggest an age after which a work is unlikely to have a rightsholder that would object to reproduction for a patron request, or parameters by which an archive might decide on an age cutoff to use for its patron-driven digitization requests. By following a scalable workflow like the one described in this article, archive staff could respond to patron requests for remote access without staff needing to do time-consuming research involving unpredictable timelines, thereby responding to researchers in a timely fashion. While not all items requested by users could be digitized under this

³ In this article we have not established a brightline for how old is "sufficiently old." This is one of the matters to be resolved through further research, as discussed in the following section.

model, many could, thereby increasing remote access to archival collections during and even after the COVID-19 crisis.



References

- American Association of Law Libraries. (2014). *AALL guidelines on the use of copyrighted works by law libraries*. Retrieved from https://www.aallnet.org/bok_profleadership/aall-guidelines-on-the-use-of-copyrighted-works-by-law-libraries/
- Association of Research Libraries, Center for Social Media, and Program on Information Justice and Intellectual Property. (2012). *Code of best practices in fair use for academic and research libraries*. Retrieved from <https://www.arl.org/wp-content/uploads/2014/01/code-of-best-practices-fair-use.pdf>
- Bailey, L., Courtney, K. K., Hansen, D., Minow, M., Schultz, J., & Wu, M. (2018, September). *Position statement on controlled digital lending*. Retrieved from Controlled Digital Lending by Libraries website: <https://controldigitallending.org/statement>
- Buttler, D. K. (2012). Intimacy gone awry: Copyright and special collections. *Journal of Library Administration*, 52(3–4), 279–293. <https://doi.org/10.1080/01930826.2012.684506>
- Cheney, M. (2020, July 6). Controlled digital lending panel recap [Blog post]. *AALL Legal Innovation and Technology SIS*. <http://blog.cssis.org/2020/07/06/controlled-digital-lending-panel-recap/>
- Copyright Act of 1976, 17 U.S.C. §§ 101–810 (1976).
- Dickson, M. (2010). Due diligence, futile effort: Copyright and the digitization of the Thomas E. Watson Papers. *American Archivist*, 73(2), 626–636. <https://doi.org/10.17723/aarc.73.2.16rh811120280434>
- Dryden, J. (2014). The role of copyright in selection for digitization. *American Archivist*, 77(1), 64–95. <https://doi.org/10.17723/aarc.77.1.3161547p1678423w>
- Evans, J., & Hernández Durán, M. (2018). Rights review for sound recordings: Strategies using risk and fair use assessments. *American Archivist*, 81(2), 323–372. <https://doi.org/10.17723/0360-9081-81.2.323>
- Fialkov, J., O’Keeffe, H., & Kaufman, C. D. (2017, December 2). *Put it on your bucket list: Navigating copyright to expose digital AV collections at scale*. SlideShare. https://www.slideshare.net/WGBH_Archives/put-it-on-your-bucket-list-navigating-copyright-to-expose-digital-av-collections-at-scale

- Fortney, K. (2013). Braving the present: Experience and copying risk assessment for digitizing recent historical collection. In D. M. Mueller (Ed.), *Imagine, innovate, inspire: The proceedings of the ACRL 2013 conference* (pp. 84–89). Retrieved from the Association of College and Research Libraries website: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2013/papers/Fortney_Braving.pdf
- Hirtle, P. B. (2020, January 3). *Copyright term and the public domain in the United States*. Retrieved from the Cornell University Library Copyright Information Center: <https://copyright.cornell.edu/publicdomain>
- Light, M. (2014). Managing risk with a virtual reading room: Two born digital projects. In M. Light (Ed.), *Reference and access: Innovative practices for archives and special collections* (pp. 17–35). Rowman & Littlefield.
- Public statement of library copyright specialists: Fair use & emergency remote teaching & research. (2020, March 13). GoogleDocs. Retrieved from <https://tinyurl.com/tvnty3a>
- Smith, K. L. (2012). Copyright risk management: Principles and strategies for large-scale digitization projects in special collections. *Research Library Issues: A Quarterly Report from ARL, CNI, and SPARC* (279), 17–23. <https://doi.org/10.29242/rli.279.6>
- UCLA Library Special Collections. (n.d.). Digital project toolkit.. Retrieved August 17, 2020, from the UCLA Library website: <https://www.library.ucla.edu/special-collections/programs-projects/digital-projects-special-collections>
- University of California Berkeley Library. (2020, April 28). *Responsible access workflows*. Google Slides. <https://docs.google.com/a/berkeley.edu/presentation/d/1V66PGpIq9xqXxdvngpD3rkAMoIw2hIyVVDS4Iv4VFOM/edit?usp=sharing>

