

LIBRARY OF CONGRESS COLLECTIONS POLICY STATEMENTS

SUPPLEMENTARY GUIDELINES

Open Digital Content

Contents

- I. Scope
- II. Background
- III. Diverse and Inclusive Collecting Statement
- IV. Collecting Policies
- V. Collecting Guidelines
- VI. Acquisitions Sources & Methods

Preface

The Library's traditional functions of acquiring, cataloging, preserving, and serving collection materials of historical importance to Congress and the American people extend to digital materials, including open digital content on the web and on external tangible media. While many institutions provide access to their own open digital content or link to content produced by other institutions, this document outlines the Library's policy for acquiring, archiving and making accessible open digital content produced by other institutions and individuals. As the *de facto* national library for the United States, the Library collects open digital content to address widely recognized concerns that open licenses do not always guarantee long-term preservation or access to content. The Library acquires open digital content deemed to have historical and research significance, applying the same collecting criteria used to acquire traditionally published materials. Central to these collecting efforts is the acknowledgment that open digital content is playing an increasingly important role in the intellectual, commercial, and creative life of the Library's diverse users. Given the rapidly evolving landscape of open digital content and the technologies associated with its acquisition and use, the Library will review these guidelines on a regular basis to ensure they are responsive to current and future research needs.

I. Scope

This document provides guidance for the recommending and selecting of open content in all digital formats that the Library acquires, preserves, and makes permanently accessible. For the purposes of this document, open digital content is defined as:

- any digital material that is licensed for free and open use and redistribution, such as works under a Creative Commons (CC) license,
- material that is not protected by any country's copyright law, such as U.S. government documents and public domain content,

- and copyrighted content that is available at no cost for which the Library has secured permission from the content owner to redistribute the content openly.

The Library also acquires digital content with restrictions on use and redistribution, as well as open content in analog formats; neither are covered by this statement. Guidance for these materials should be sought in the subject-specific Collections Policy Statements, in conjunction with applicable type- or format-based Collections Policy Statements and Supplementary Guidelines. Resources covered by this guidance include for example Open Access (OA) publications such as electronic serials and books; digital repositories with content uploaded by individuals or institutions; platforms with open user-created content such as Wikipedia; and any unpublished, unstructured or raw digital content on the Internet or on external tangible media.

Guidelines for recommending other digital materials such as datasets and purchased electronic resources may be found in the [Datasets Supplementary Guidelines](#) and the [Electronic Resources Supplementary Guidelines](#), respectively. Specific guidance about the Library's web archiving program as a mode of acquiring both open and restricted websites is available in the [Web Archiving Supplementary Guidelines](#). Subject-focused collecting guidance for U.S. and foreign government publications is available in the Collections Policy Statements for [Government Publications – United States](#) and [Government Publications – Non-U.S.](#) When considering open content for acquisition, the Open Digital Content Supplementary Guidelines should be used in conjunction with the subject-focused [Collections Policy Statements](#), which must be consulted to determine whether resources are in scope.

II. Background

The Library has traditionally prioritized no-cost acquisitions in all subject areas through its gift and exchange programs and has sought to share its collections as broadly as permissions allow. The Library has a specific research strength in open content acquired through its web archiving program, which has been used to actively collect in a wide range of subject areas since 2000. The Library of Congress Web Archive Collections contain a vast body of open content for which permissions were explicitly solicited by the Library or for which open licenses and broad redistribution terms were identifiable, allowing the Library to make this content widely available to researchers. The Web Archive Collections range from unique cultural material such as repositories of memes, emojis, and web comics to select websites of U.S. and foreign government agencies, U.S. House and Senate offices and committees, and U.S. and foreign elections. Other examples of notable open content resources collected by the Library include English-language content from the free online encyclopedia Wikipedia, peer-reviewed academic open access e-books from the Directory of Open Access Books (DOAB), serials deposited with the Copyright Office by the Open Access (OA) publisher Hindawi, select foreign government publications no longer available in print, public domain sound recordings in the National Jukebox, free to use and reuse motion pictures in the National Screening Room, and born-digital or digitized prints and photographs with no rights restrictions including the official portraits of contemporary U.S. Presidents transferred by the White House.

III. Diverse and Inclusive Collecting Statement

As the nation's de facto national library, the Library of Congress strives to build an expansive, yet

selective, collection that records the creativity of the United States and is reflective of the nation's diversity and complexity. The Library's mandate is to have collections that are inclusive and representative of a diversity of creators and ideas. A priority includes acquiring material of underrepresented perspectives and voices in the Library's collections to ensure diverse authorship, points of view, cultural identities, and other historical or cultural factors. The Library also seeks to build a research collection that comprises a globally representative sample of international materials that are diverse in voice and perspective, relative to their places of origin, further supporting the Library's mission to sustain and preserve a universal collection of knowledge and creativity for Congress and future generations.

Diverse collecting is mentioned within many of the Library's Collections Policy Statements. In addition, the Library has adopted several specific collection policies in an effort to ensure it is building an inclusive and representative collection. For more information, see the Library's Collections Policy Statements on [Ethnic Materials](#), [LGBTQIA+ Studies](#), [Women's and Gender Studies](#), [Independently Published and Self-Published Textual Materials](#), and [Countries and Regions with Acquisitions Challenges](#).

IV. Collecting Policies

The Library seeks to selectively acquire a diverse collection of digital open content. Its collecting in this area should reflect the breadth and depth of the materials available, particularly in areas of publication or creation for which there is no analog surrogate. In terms of subject matter, open content acquired for the Library's collections is guided by the relevant [Collections Policy Statements](#).

The Library generally acquires open content for which terms allow local hosting and wide redistribution of content. However, in select cases when digital content is deemed by Recommending Officers to be endangered and of high cultural or research value, the Library may acquire the material before the full terms allowing the Library to acquire and redistribute content are known or before permission to acquire and redistribute rights-restricted content has been solicited.

The Library will not acquire any open digital content for which there is an indefinite embargo on public access preventing the Library from making the content accessible in the future.

V. Collecting Guidelines

The Library of Congress will acquire freely available and redistributable open content resources in various formats. In particular, Recommending Officers should note the following guidelines when evaluating resources:

- The acquisition and redistribution terms for content should be known before an item is acquired.
 - Does the material have an open license or explicit open use and distribution terms listed on the work?
 - If not, are there use and distribution terms listed on the platform or site providing the material? Examples: open access (OA) scholarly literature, CC licensed works, open software with licenses such as GNU, and other similarly licensed works that explicitly allow for acquisition (local hosting) and wide redistribution (serve onsite AND offsite via

loc.gov). Does the site where the material of interest is found belong to the content owner? If not, how has it been determined that the material is open content?

- For individual items that do not have an open license or traditional copyright terms listed, Library staff will review the website, platform, or database-level use and distribution terms to determine the terms and conditions that apply. Conversely, for individual items that have an open license or open use terms listed but for which platform-level terms may be more restrictive, additional legal guidance may be required.
- If a specific license or explicit terms are not identifiable for a work, any available contact information may be used to attempt to acquire the content via web archiving, using the Copyright Information and Permissions form, or other methods of acquisition.
- Preference is for resources that do not require user registration in order to download or gain access to content. However, exceptions may be made in the case of digital content that is endangered and of high cultural or research value, provided that the Library has secured permission from the content owner for open redistribution.
- In general, U.S. government works are open content, but occasionally a federal publication may include a portion of content from a third party who holds the copyright. States, territories, and municipalities also may copyright their publications. Additionally, classified government works, including controlled “For Official Use Only” works, are not redistributable and thus out of scope for open content collecting. For more guidance on U.S. government works, see the [Government Publications – United States](#) Collections Policy Statement.
- The Library’s [Recommended Formats Statement](#) should be considered when identifying open content for acquisition.
 - It is preferred that the format of open content meet the specifications for Preferred or Acceptable formats outlined in the Statement.
 - Of particular relevance are specifications for digital textual works such as e-books and e-serials; websites; still image works/digital photographs; audio works, and moving image works.
- The scale and extent (size and versions) should be determined when evaluating a resource.
- Selection and Recommending Officers evaluating open content are encouraged to be aware of the highly interdisciplinary nature of many digital resources in this genre and to consult with colleagues when judging the appropriateness of open content material for the Library’s collections.

VI. Acquisitions Sources & Methods

The primary means by which the Library has built its collections of digital open content are the web archiving program and gifts. The Copyright Acquisition Division’s mandatory and voluntary deposit programs represent another developing acquisitions pathway. Other means of acquisition, e.g., exchange and government transfer, also have contributed to the strength of the Library’s open content collections. As the nature of open content changes, so too will the Library’s methods for acquiring content, which currently include web archiving, direct download, shipment of digital files on hard drives, secure file transfer, and API-based accessioning.

The Library acquires diverse genres and forms of open content. However, the following broad sources are particularly important in this area of collecting:

- Public domain content, which is preferred over rights-restricted open content
- Bulk sets of openly available user-generated creative works in all digital formats
- Open access e-journals and e-books
- Open educational resources (OERs)
- Government publications
- Digital versions of continuing print resources previously collected by the Library that are no longer available in print or are preferred in digital format

August 2020. Updated by the Collection Development Office July 2022.