

# LIBRARY OF CONGRESS COLLECTIONS POLICY STATEMENTS

## SUPPLEMENTARY GUIDELINES

### Social Media

#### Contents

- I. Scope
- II. Diverse and Inclusive Collecting Statement
- III. Background
  - A. Collecting Social Media at the Library of Congress
  - B. Social Media Definitions and Examples
- IV. Collecting Policy
- V. Exceptions

#### I. Scope

The purpose of this document is to provide guidance regarding the range of social media content that may be collected by the Library of Congress as the *de facto* national library of the United States, and to indicate the types of such content that are not to be collected. Prior to this policy, there was no formal guidance. A large amount of social media content has been collected via web harvesting, and it currently is maintained in the Library's web archives. In addition, the Library has a large Twitter Archive, which was acquired over several years via a technical process different from web harvesting. That collection is under embargo for an indefinite period. This new policy pertains to future collecting and is not to be retrospectively applied to content that has already been acquired.

The guidance provided in this document applies to the collecting of digital social media content regardless of the means of acquisition, including but not limited to web harvesting and acquisition of bulk data. For additional guidance, refer to the Supplementary Guidelines for [Web Archiving](#) and for [Datasets](#).

Most of this document pertains to collecting via web harvesting. Collecting via other modes of acquisition (for example, a gift of content from a social media platform) is not yet common. Until this policy is updated at a future time, the appropriate Library staff member (a Recommending Officer) should bring such proposed acquisitions to the attention of the Collection Development Officer, who will consult as necessary and make a determination on the appropriateness of the acquisition.

#### II. 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](#).

### III. Background

#### A. Collecting Social Media at the Library of Congress

The Library has been collecting social media content for years. For example, the web archives include captures of Facebook dating back over a decade. Social media comprises a significant portion of the Library's web archives, and is primarily collected via its connection to other web pages. That is, social media sites are generally collected based on their association with a website that is being acquired via the Library's web archiving program. As with all web archiving conducted by the Library of Congress, social media web archiving is impacted by the Library's permissions process. Under this process, the minimum permissions level indicates that notice must be provided to the social media page owner, with the exception of U.S. government pages or those that use Creative Commons or similar terms of service. Refer to [the Library's webpage about the Web Archiving program](#) for more information.

Web archiving is the most common mode of acquisition for social media content at the Library. However, social media content is distinct in that it has been collected by both web archiving and file transfer. The Library's Twitter Archive, a closed collection that was acquired by receiving tweet text directly from Twitter, is an example of the latter mode of acquisition. The Library continues to preserve and secure the Twitter Archive, which remains embargoed until access issues can be resolved in a cost-effective and sustainable manner. The Library no longer directly acquires tweet text from Twitter, but intends to use the web archiving program to selectively acquire Twitter content in scope for the [Collections Policy Statements](#).

Note that due to technical difficulties that apply across the evolving landscape of social media web archiving, at the time of writing this Supplementary Guidelines document, much social media content, including Twitter, Facebook, Instagram, and YouTube, cannot currently be captured and playback of older captures is slow. The Library's collecting goals and guidelines in this document continue to apply as this industry-wide issue with content capture and playback remains.

#### B. Social Media Definitions and Examples

The term *social media* refers to certain internet-based services and platforms that enable communication within groups of users. The communication that is enabled by these and other platforms is referred to as *social networking*. The act of social networking happens in venues across the internet - not just via services considered to be social media. For instance, online comments to a posted article on a website are considered to be social networking. This policy is intended to apply to social networking content from any platform, including standard websites.

Such services share some common features including the following.<sup>1</sup>

- “User-generated content is the lifeblood of social media.
- Individuals and groups create user-specific profiles for a site or app designed and maintained by a social media service.
- Social media services facilitate the development of social networks online by connecting a profile with those of other individuals and/or groups.”

User-generated content can include material such as text posts or comments, digital photos or videos, and data generated through all online interactions.

Major social media platforms include:

- Facebook
- Twitter
- Instagram
- YouTube
- Pinterest
- LinkedIn
- Reddit
- TikTok
- Snapchat

#### IV. Collecting Policy

The following types of collecting are allowed, as guided by the relevant [Collections Policy Statements](#), except where specifically prohibited later in this section.

- Social media accounts associated with a website that is being acquired via the Library’s web archiving program may also be collected in accordance with the Library’s country/category permissions structure for the associated website.
- Standalone social media accounts of individuals or groups may be collected, in accordance with the Library’s country/category permissions structure for the web archiving program, including the current notice/notice approach for United States election candidate sites.

The collecting of the following types of content is not allowed.

- Social media content that is collected by hashtag or similar linked subject tags (e.g., Twitter hashtags), or social media content posted in a shared feed, group, forum, community, or similar mechanism by multiple users (e.g., Reddit). This is referred to as *Collective Social Media*. It is created by various authors, thus making the securing of permissions difficult or impossible.
- Any content that cannot eventually be made accessible to the Library’s users (for technical or legal reasons, etc.).

---

<sup>1</sup> Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, 745-750. Retrieved from <https://doi.org/10.1016/j.telpol.2015.07.014>

## V. Exceptions

If a fair use justification can be made for acquiring *Collective Social Media*, as in the case of election-related content, the appropriate staff member (a Recommending Officer) should send a formal request to the Collection Development Officer. Included in the request should be the following information, at a minimum: 1) the associated web archives collection, 2) the web archiving category, 3) specific links to the content to be collected, and 4) a brief written justification for the request.

The Collection Development Officer will then consult with the Office of the General Counsel and Library senior managers as appropriate to determine whether an exception to the general policy will be approved.

November 2021. Updated by the Collection Development Office July 2022.