



Archive-It



What is it? Archive-It is a community of partner organizations and their shared, "end-to-end" web archiving service, administered by the Internet Archive (archive.org).

Who should use it? Archivists looking for technical and organizational partners to help them through the entire web archiving process, from collection development through long-term preservation, and/or looking to archive web content of multiple types and sizes.

Key benefits: User-friendly web application; Dedicated technical support and training; Full-text search and descriptive metadata features; Built-in access layer and "Wayback Machine"-style replay.

What's the catch? For the human and engineering resources that it devotes to upkeep and storage, the Internet Archive requires most partners to provide an annual subscription fee.

Online: archive-it.org

Save Page Now



What is it? Save Page Now is a tool for quickly (and freely) adding one web page at-a-time to the Internet Archive's web-spanning "Wayback Machine." Just enter a URL into the box in order to archive it for immediate storage and replay.

Who should use it? Archivists interested in pointing to single-page items in an archival, rather than live, forms. Using Save Page Now is the fastest way to get a "permalink" for any web resource that you want to share/promote in a form safe from possible deprecation.

Key benefits: Free to use; Immediate, on-demand archiving; No web expertise required.

What's the catch? While exceptionally fast and easy to use, Save Page Now cannot archive all web pages. Webmasters and site owners can prevent this tool from archiving some or all of their content with simple technical tweaks on their end.

Online: archive.org/web

Social Feed Manager

Social Feed Manager

What is it? Social Feed Manager is a software developed by the George Washington University Libraries in order to collect web data specific to social media services like Twitter and Instagram.

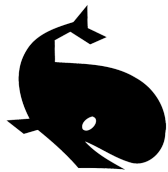
Who should use it? Technically adept archivists who wish to collect select data from social media services, accounts, hashtags, etc., in forms conducive to large-scale data analysis and visualization.

Key benefits: Integration of data APIs across multiple social media platforms; Dedicated community of adopters and open source developers.

What's the catch? Installing, using, and optimizing Social Feed Manager for your institution's collecting scope requires some confidence in computer programming from the command line.

Online: gwu-libraries.github.io/sfm-ui

WAIL



What is it? WAIL (Web Archiving Integration Layer) is an open source and free-to-use graphical user interface (GUI) that enables users to run the same kinds of capture and replay technologies that the Internet Archive and other web archivists use from their own machines, but without regular, advanced use from the command line interface once it is installed.

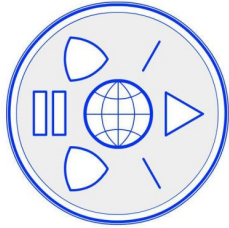
Who should use it? Archivists who wish to collect large but technically simple (more 'static' than 'dynamic') websites on an on-demand basis for preservation in their own local digital repository and replay through their own dedicated web server.

Key benefits: "One-click preservation"; Reliance on strictly open source component software.

What's the catch? Lacking the enhancements that services like Archive-It and Webrecorder offer for archiving dynamic web content, WAIL cannot be used to archive the more interactive sites and social media feeds common to the modern web. And with its very particular capabilities and dependencies on local support, it has a small user base and community of contributors/developers.

Online: machawk1.github.io/wail

Webrecorder



What is it? Webrecorder is an easy-to-use tool that archives web content as it is viewed in web browser by a user, generates archival files for download and local storage/preservation, and provides a simple access layer for immediately sharing archived content online. It is especially capable at “recording” user-responsive and interactive material that can be challenging to automated acquisition tools, which is why it is used often in the conservation of net art; since 2015 it has been a project of the online arts organization Rhizome.

Who should use it? Archivists archiving highly dynamic and/or experimental web content, such as artists’ sites or web-based works.

Key benefits: Free and easy to use; Built-in access layer with some descriptive features; High fidelity online replay from many traditionally challenging media types.

What’s the catch? Scale: Archiving web content through manual browsing becomes increasingly laborious as collecting scopes grow larger, and while there is no limit to the amount that may be archived and downloaded to local machines, Webrecorder users’ online-accessible collections are capped at 5GB of total data.

Online: webrecorder.io

youtube-dl



What is it? youtube-dl is a command line tool for harvesting video from YouTube and similar audio and video streaming platforms, like Vimeo, Soundcloud, and more. It enables users to download and store local copies of these media files in their own repositories.

Who should use it? Technically proficient archivists who wish to collect a large amount of online audio or video content related to their institution, an event, person, etc., for local management and preservation, and without investing in a wider web archiving program.

Key benefits: Free to use; Open source tool with broad user base of active developers.

What’s the catch? Usage beyond the most basic type requires confidence in command line programming and rigorous quality assurance, as web-based audio and video platforms change frequently and significantly, and the open source community catches up. No functionalities beyond data capture.

Online: rg3.github.io/youtube-dl