

Help:Using the Wayback Machine



This help page is a how-to guide.

It details processes or procedures of some aspect(s) of Wikipedia's norms and practices. It is not one of Wikipedia's policies or guidelines, and may reflect varying levels of consensus and vetting.

The Wayback Machine is a service which can be used to cite archived copies of web pages used by articles. This is useful if a web page has changed, moved, or disappeared; links to the original content can be retained. This process can be performed automatically, using the web interface for User:InternetArchiveBot.

Editors are encouraged to add an archive link as a part of each citation, or at least submit the referenced URL for archiving, at the same time that each citation is created or updated. New URLs added to Wikipedia articles (but not other pages) are usually automatically archived by a bot.

Visit the webform at <https://web.archive.org>, enter the original URL of the web page of interest in the "Wayback Machine" search box and then hit return/enter. The next screen may:

- show a calendar listing the snapshot dates for all archived copies of that page, or
- show a box near the bottom of the page with a link inviting the user to **Save this url in the Wayback Machine**,

In short, this is the code that needs to be added to an existing `{{cite web}}` or similar template:

```
<ref>{{cite ... <!--EXISTING REFERENCE--> |archive-url=https://web.archive.org/web/<date>/http://www.originalurl.com
|archive-date=<date> |url-status=dead}}</ref>
```

Contents

URL formats

[Initial request](#)

[Specific archive copy](#)

[Removing the navigational toolbar](#)

[Latest archive copy](#)

Limitations

JavaScript bookmarklet

[To see a dead page](#)

[To save a live page](#)

Command-line tool

[Save a live page](#)

[Get the oldest archive](#)

[Get the newest archive](#)

[Archive close to a date and time](#)

Browser add-ons and apps

Using the webarchive template

See also

[Docs](#)

[Tools](#)

References

URL formats

A link to the Wayback Machine usually starts with `https://web.archive.org/web/` followed either by a single asterisk or a 14-digit datetime reference, then a slash and finally the URL of the original web page.

Initial request

The following example requests archived copies of the main index page of Wikipedia. Such requests usually result in a calendar with links to all archived copies of the requested page.

- https://web.archive.org/web/*/http://en.wikipedia.org/ (https://web.archive.org/web/*/http://en.wikipedia.org/)

Use the above URL format to discover the extent to which the requested page has been archived. Click one of the highlighted dates to select that specific archived copy.

It is possible to narrow down the request by providing a date code with fewer than 14 digits followed by * (in this example, display only archived snapshots matching December 2005)

- https://web.archive.org/web/200512*/http://www.wikipedia.org/ (https://web.archive.org/web/200512*/http://www.wikipedia.org/)

If the target web page hasn't yet been archived, a box appears near the bottom of the page with a link inviting the user to **Save this url in the Wayback Machine**. Clicking this invokes a request to

- <https://web.archive.org/save/http://www.wikipedia.org/> (<https://web.archive.org/save/http://www.wikipedia.org/>)

The above URL will show the current version of the requested web page and start the process that will attempt to archive the web page. If successful, the archived copy will become available immediately after the process is completed.

For some requested pages, the Wayback Machine will return an error message explaining why that particular page has not and cannot be archived. In those cases, try a different archiving service such as WebCite.

Specific archive copy

Once the target web page has been archived, each of the specific dated archives can be individually requested using the format shown below.

The next example links to the archived copy of the main index page of Wikipedia exactly as it appeared on 30 September 2002 at 12:35:25 pm in the UTC timezone. The datetime format is `YYYYMMDDhhmmss`.

- <https://web.archive.org/web/20020930123525/http://www.wikipedia.org/> (<https://web.archive.org/web/20020930123525/http://www.wikipedia.org/>)

Use the above format to link directly to a specific archive copy.

Adding an asterisk immediately after the date (or in place of it) is a quick way to show the calendar view of all archived copies.

The following flags can be appended to the datetime field to modify the format in which the archived content is displayed^{[1][2]}:

- **id_** Identity - perform no alterations of the original resource, return it as it was archived.
- **js_** JavaScript - return document marked up as JavaScript.
- **cs_** CSS - return document marked up as CSS.
- **im_** Image - return document as an image.
- **if_** or **fw_** Iframe - return document formatted normally, but without the navigational toolbar.

Depending on the circumstances under which the page images were archived, the rendering of these pages may not be consistent; therefore, it is recommended that the flags be tested before being incorporated into Wikipedia documents. The datetime format is YYYYMMDDhhmmss, followed by an optional formatting flag, such as the ones above.

Removing the navigational toolbar

Normally, when displaying an archived web page, the Wayback Machine will rewrite parts of the underlying code (such as CSS/image references), in order to make the page look as similar as possible to how it looked at the time the page was archived. By default, it will also add a navigational toolbar. This toolbar is undesirable for links to a specific known archived copy of the page.

The **id_** "identity" flag was previously recommended to return the page *exactly* as it was archived, without the toolbar. Unfortunately, many pages will render poorly with this flag because the CSS/image references are not fixed to use archived copies of those resources.

A better choice is the **if_** "iframe" flag, which omits the toolbar while still fixing the references. This will make the rendered page look as similar to the original web page as possible.

For example, here is an archived post discussing the **id_** identity flag. This is a normal link to the Wayback Machine, which renders **with** the navigational toolbar:

- <https://web.archive.org/web/20130806040521/http://faq.web.archive.org/page-without-wayback-code/> (<https://web.archive.org/web/20130806040521/http://faq.web.archive.org/page-without-wayback-code/>)

Here is the same archived page, with the **id_** identity flag added to the link. This does not include the toolbar, but now the page renders poorly because of the broken references:

- https://web.archive.org/web/20130806040521id_/http://faq.web.archive.org/page-without-wayback-code/ (https://web.archive.org/web/20130806040521id_/http://faq.web.archive.org/page-without-wayback-code/)

Finally, here is the same archived page, with the **if_** iframe flag instead. This renders perfectly, without the toolbar:

- https://web.archive.org/web/20130806040521if_/http://faq.web.archive.org/page-without-wayback-code/ (https://web.archive.org/web/20130806040521if_/http://faq.web.archive.org/page-without-wayback-code/)

Since this is the most faithful reproduction of the original web page, please use the **if_** iframe flag for links to specific archive copies!

Latest archive copy

The next example links to the most current version of the archived page.

- <https://web.archive.org/web/http://www.wikipedia.org/> (<https://web.archive.org/web/http://www.wikipedia.org:80/>)

Using the above format is discouraged. The request is redirected to the long-form URL, including a 14-digit datetime stamp, for the latest archive copy thereby defeating the purpose of using the archive to link directly to a specific old version of the page.

Likewise, a similar archive URL but with the number **1000** links to the oldest archive copy.

- <https://web.archive.org/web/1000/http://www.wikipedia.org/> (<https://web.archive.org/web/1000/https://www.wikipedia.org/>)

See also: [Advanced URL locator hints and tips \(https://archive.org/web/web-advancedsearch.php\)](https://archive.org/web/web-advancedsearch.php) – Internet Archive

Limitations

Before October 2013, it would often take weeks or months for an archived copy of a web page to become available. Nowadays, a request to archive a particular web page is actioned immediately and the result is usually made available within minutes.

Prior to April 2017,^[3] The Internet Archive honored the [robots exclusion standard](#). It would not archive sites that disallow access, and it would remove access to previous versions of a disallowed page.

For example, *The New York Times* once had a robots.txt page at <https://www.nytimes.com/robots.txt> which included:

```
User-agent: *
Disallow: /online/
Disallow: /archives/
Disallow: /reuters/
```

Thus, archive requests for URLs within those folders of *The New York Times*'s website would be rejected.

JavaScript bookmarklet

A [bookmarklet](#) is a one-click button in a web browser that is stored like a bookmark but uses javascript to carry out certain actions.

To see a dead page

To use a bookmarklet when you're at a [dead link](#) web page and want to visit archives saved by the [Wayback Machine](#), click and drag the following code to your browser's bookmarks toolbar, then name it something memorable, such as **Wayback** (e.g. [Wayback \(http://javascript:void\(window.open\('https://web.archive.org/web/*/'+location.href\)\)](http://javascript:void(window.open('https://web.archive.org/web/*/'+location.href))))):

```
javascript:void(window.open('https://web.archive.org/web/*/'+location.href.replace(/\$/ , ' ')));
```

Then, when you are at a dead page, you may click the bookmarklet and it will automatically take you to the Wayback Machine's archives of that page.

The preceding code may not work for all users. In that case, you may try the following bookmarklet:

```
javascript:location.href='https://web.archive.org/web/*/'+document.location.href.replace(/\$/ , '');
```

To save a live page

For a bookmarklet that allows you to manually archive a page you are visiting, store the following code in a bookmark on your browser's toolbar, with a name such as **Wayback Save** (e.g. [Wayback Save \(http://javascript:void\(window.open\('https://web.archive.org/save/'+location.href\)\)\)](http://javascript:void(window.open('https://web.archive.org/save/'+location.href)))):

```
javascript:void(window.open('https://web.archive.org/save/'+location.href));
```

Command-line tool

Install [waybackpy](https://github.com/akamhy/waybackpy) (<https://github.com/akamhy/waybackpy>) - ([PyPi \(https://pypi.org/project/waybackpy/\)](https://pypi.org/project/waybackpy/)) - ([docker \(https://hub.docker.com/r/secsi/waybackpy\)](https://hub.docker.com/r/secsi/waybackpy)) - ([snapcraft.io \(https://snapcraft.io/waybackpy\)](https://snapcraft.io/waybackpy))

Waybackpy is an OS-independent [command-line tool](#) and a [Python package](#) that interfaces the [Internet Archive's Wayback Machine APIs](#) (Save API , Availability API and CDX API).

Save a live page

```
waybackpy --url "https://en.wikipedia.org/wiki/Social_media" --save
```

```
Archive URL:
https://web.archive.org/web/20220101114012/https://en.wikipedia.org/wiki/Social_media
Cached save:
False
```

The line below 'Archive URL:' contains the archive URL and the line below 'Cached save:' indicates whether the URL returned by Wayback Machine was saved before the request was made by the client, thus cached save.

Get the oldest archive

waybackpy uses the Wayback Machine's CDX Server API for retrieving the oldest archive.

```
waybackpy --url "https://en.wikipedia.org/wiki/SpaceX" --oldest
```

```
Archive URL:
https://web.archive.org/web/20040803000845/http://en.wikipedia.org:80/wiki/SpaceX
```

Get the newest archive

```
waybackpy --url "https://en.wikipedia.org/wiki/YouTube" --newest
```

```
Archive URL:
https://web.archive.org/web/20211207121553/https://en.wikipedia.org/wiki/YouTube
```

Archive close to a date and time

To find the archive of google.com close to 2008-08-08 08:08 UTC (8th of August, 2008 and 8 minutes past the 8th hour UTC time) use the following command. You may omit the flags you don't care about. Wayback Machine timestamps are UTC time.

```
waybackpy --url google.com --near --year 2008 --month 8 --day 8 --hour 8 --minute 8
```

```
Archive URL:  
https://web.archive.org/web/20080808051143/http://google.com/
```

Browser add-ons and apps

The Internet Archive provides a browser add-on that can be used to easily access pages on the Wayback Machine for the currently viewed site, along with options to save a copy of the page to the Wayback Machine. Currently, versions of the add-on are available for Google Chrome (<https://chrome.google.com/webstore/detail/wayback-machine/fpnmgdkabkmnadcjpehmlllknepkmiak>), Mozilla Firefox (https://addons.mozilla.org/en-US/firefox/addon/wayback-machine_new/), and Safari (<https://apps.apple.com/app/wayback-machine/id1472432422>).

Additionally, apps for iOS (<https://itunes.apple.com/us/app/wayback-machine/id1201888313>) and Android (<https://play.google.com/store/apps/details?id=com.archive.waybackmachine>) are available for mobile devices.

Using the webarchive template

{{webarchive}} is an easy way to create very basic links to the Wayback Machine (or other archiving services). It typically isn't used for citations since it doesn't include information like author, date, and publication, but it can be useful for non-citation links. Use the `|url=`, `|title=` and `|date=` parameters to specify the URL, title and archive date. For example:

- {{webarchive |url=https://web.archive.org/web/20010727112808/http://www.wikipedia.org/ |date=July 27, 2001 |title=Wikipedia }}
→ Wikipedia (<https://web.archive.org/web/20010727112808/http://www.wikipedia.org/>) at the Wayback Machine (archived July 27, 2001)

Without the date included:

- {{webarchive |url=https://web.archive.org/web/*/http://www.wikipedia.org/ |date=* |title=Wikipedia }}
→ Wikipedia (https://web.archive.org/web/*/http://www.wikipedia.org/) at the Wayback Machine (archive index)

See the {{webarchive}} documentation for additional options

Working with cite templates

{{citation}}, and all of the Citation Style 1 templates support the `|archive-url=` parameter (Note that the `|archive-date=` parameter is also required). Other citation templates may also support `|archive-url=` — see their documentation.

- {{citation |url=http://www.wikipedia.org/ |title=Wikipedia Main Page |archive-url=https://web.archive.org/web/20020930123525/http://www.wikipedia.org/ |archive-date=2002-09-30 |access-date=2005-07-06 }}

→ "Wikipedia Main Page" (<https://web.archive.org/web/20020930123525/http://www.wikipedia.org/>). Archived from the original (<https://www.wikipedia.org/>) on 2002-09-30. Retrieved 2005-07-06.

- Where an archived resource notes its original publication date, use `|date=` in place of `|access-date=`.
- When adding an archive URL to any citation where the original resource URL is still working, it is useful to add the `|url-status=live` parameter. With `|url-status=live`, clicking the title in the footnote invokes the original (live) URL, clicking "Archived" gives the archived copy. Otherwise the title invokes the archived page, "Original" invokes the (dead unless it has been reinstated) original link:

```
{{citation |url=http://www.wikipedia.org/ |title=Wikipedia Main Page |archive-url=https://web.archive.org/web/20020930123525/http://www.wikipedia.org/ |archive-date=2002-09-30 |access-date=2005-07-06 |url-status=live }}
```

→ "Wikipedia Main Page" (<https://www.wikipedia.org/>). Archived (<https://web.archive.org/web/20020930123525/http://www.wikipedia.org/>) from the original on 2002-09-30. Retrieved 2005-07-06.
Should the original URL stop working, it is a simple job to either change this to `|url-status=dead` or remove the parameter.

See also

- [Help:Archiving a source](#)

Docs

- [Wikipedia:Link rot](#), how-to guide for prevention of link rot
- [Using WebCite](#), how-to guide
- [Using Archive.is](#), how-to guide
- [Using Perma.cc](#), how-to

Tools

- `{{dead link}}`, for flagging dead links
- `{{linkrot}}`, for flagging pages with bare links
- `{{user archiveurl}}`, userbox
- `{{user Internet Archive}}`, userbox
- `{{user web archive}}`, userbox

References

1. "Wayback Administrator Manual" (http://archive-access.sourceforge.net/projects/wayback/administrator_manual.html). Internet Archive. Archived (https://web.archive.org/web/20140120095314/http://archive-access.sourceforge.net/projects/wayback/administrator_manual.html) from the original on 2014-01-20.
 2. "How can I view a page without the Wayback code in it?" (<https://web.archive.org/web/20130806040521/http://faq.web.archive.org/page-without-wayback-code/>). Internet Archive. Archived from the original (<http://faq.web.archive.org/page-without-wayback-code/>) on 2013-08-06.
 3. "Internet Archive will ignore robots.txt files to keep historical record accurate" (<https://www.digitaltrends.com/computing/internet-archive-robots-txt/#ixzz4gQYOqpUi>). *Digital Trends*. 2017-04-24. Retrieved 2018-05-20.
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Help:Using_the_Wayback_Machine&oldid=1082953432"

This page was last edited on 16 April 2022, at 03:31 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License 3.0; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit

organization.