archive.today

archive.today (or **archive.ph** or **archive.is**) is a web archiving site, founded in 2012, that saves snapshots on demand, and has support for JavaScript-heavy sites such as Google Maps and progressive web apps such as Twitter.[4] archive.today records two snapshots: one replicates the original webpage including any functional live links; the other is a screenshot of the page.[5]

Features

Functionality

archive.today can capture individual pages in response to explicit user requests.[6][7][8] Since its beginning, archive.today has supported crawling pages with URLs containing the now-deprecated hash-bang fragment (#!).[9]
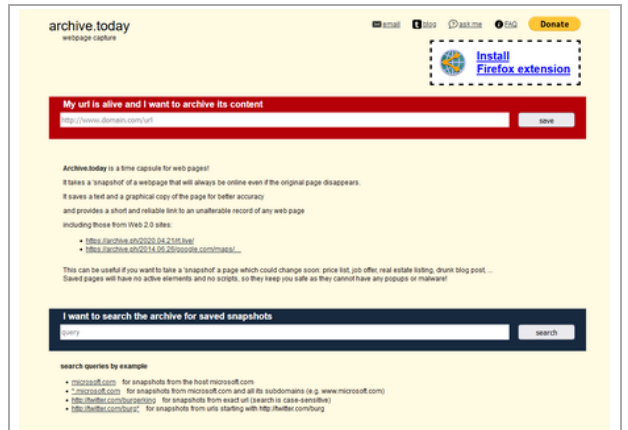
archive.today records only text and images, excluding XML, RTF, spreadsheet (xls or ods) and other non-static content. However, videos for certain sites, like Twitter, are saved.[10] It keeps track of the history of snapshots saved, requesting confirmation before adding a new snapshot of an already saved page.[11][12]

Pages are captured at a browser width of 1,024 pixels. CSS is converted to inline CSS, removing responsive web design and selectors such as :hover and :active. Content generated using JavaScript during the crawling process appears in a frozen state.[13] HTML class names are preserved inside the old-class attribute. When text is selected, a JavaScript applet generates a URL fragment seen in the browser's address bar that automatically highlights that portion of the text when visited again.

Web pages cannot be duplicated from archive.today to web.archive.org as second-level backup, as archive.today places an exclusion for Wayback Machine and does not save its snapshots in WARC format. The reverse—from web.archive.org to archive.today—is possible,[14] but the copy usually takes more time than a direct capture. Some web sites get deleted from Internet Archive's listings retroactively or blocked from being saved due to their robots.txt file, but archive.today does not use this.[8]

The research toolbar enables advanced keywords operators, using * as the wildcard character. A couple of quotation marks address the search to an exact sequence of keywords present in the title or in the body of the webpage, whereas the *insite* operator restricts it to a specific Internet domain.[15]

---

archive.today

Screenshot of the archive.today home page

Type of site	Web archiving
Available in	Multilingual
URL	archive.today (https://archive.today) (main)
	archive.ph (https://archive.ph)
	archive.is (https://archive.is)
	archive.li (https://archive.li)
	archive.vn (https://archive.vn)
	archive.fo (https://archive.fo)
	archive.md (https://archive.md)
	archivecaslytosk.onion 🧅 (Accessing link help)
	archiveiya74codqgiixo33q62qlrqtkg

Once a web page is archived, it cannot be deleted directly by any Internet user.[16] Removing advertisements, popups or expanding links from archived pages is possible by asking the owner to do it on his blog.[17]

While saving a dynamic list, archive.today searchbox shows only a result that links the previous and the following section of the list (e.g. 20 links for page).[18] The other web pages saved are filtered, and sometimes may be found by one of their occurrences.[19]

mcitqx5u2oeqnm n5bpcbiyd.onion Tór(Accessing link help)[1]	
Registration	No
Launched	May 16, 2012[2][3]

The search feature is backed by Google CustomSearch. If it delivers no results, archive.today attempts to utilize Yandex Search.[20]

While saving a page, a list of URLs for individual page elements and their content sizes, HTTP statuses and MIME types is shown. This list can only be viewed during the crawling process.

One can download archived pages as a ZIP file, except pages archived since 29 November 2019, when archive.today changed their browser engine from PhantomJS to Chromium.[21]

Since July 2013, archive.today supports the API of the Memento Project.[22][23]

History

archive.today was founded in 2012. The site originally branded itself as archive.today, but in May 2015, changed the primary mirror to archive.is.[24]

In January 2019, it began to deprecate the archive.is domain in favor of the archive.today mirror.[25]

Worldwide availability

Australia

In March 2019, the site was blocked for six months by several Australian internet providers in the aftermath of the Christchurch mosque shootings in an attempt to limit distribution of the footage of the attack.[26][27] It has since been unblocked.

China

According to GreatFire.org, archive.today has been blocked in China since March 2016,[28] archive.li since September 2017,[29] archive.fo since July 2018,[30] as well as archive.ph since December 2019.[31]

Finland

On 21 July 2015, the operators blocked access to the service from all Finnish IP addresses, stating on Twitter that they did this in order to avoid escalating a dispute they allegedly had with the Finnish government.[32] It has since been unblocked.

Russia

In Russia, only HTTP access is possible; HTTPS connections are blocked.[33][34]

Cloudflare DNS availability

As of May 2018, it has not been possible to reach the site when using Cloudflare's 1.1.1.1 DNS service.[35]

Cloudflare staff have stated that the problem is on the end of archive.today, as its authoritative nameservers return invalid records when queried from within Cloudflare's network because archive.today returns invalid data to DNS requests coming via Cloudflare's DNS servers.

archive.today's reasoning for this is the fact Cloudflare does not send EDNS Client Subnet information in its DNS requests.[36][37]

The site is available again when using Cloudflare's DNS since May 2022.[38]

See also

- Digital preservation
- List of Web archiving initiatives
- Link rot
- Perma.cc
- Wayback Machine
- Web archiving
- WebCite

Internet portal

References

1. @archiveis (30 October 2019). "a current list of all tor domains and clear net domains" (https://twitter.com/archiveis/status/1189322374598053890) (Tweet) – via Twitter.
2. Archive.is blog—When did the Archive-is site originally launch? (https://archive.today/2021.03.20-075425/https://blog.archive.today/post/77015559437/when-did-the-archive-is-site-originally-launch) at archive.today (archived 20 March 2021)
3. Archive.is — Викиреальность (https://archive.today/IfKji) at archive.today (archived 29 April 2021)
4. Brinkmann, Martin (22 April 2015). "Create publicly available web page archives with Archive.is" (https://www.ghacks.net/2015/04/22/create-publicly-available-web-page-archives-with-archive-is/). *Ghacks*. Archived (https://web.archive.org/web/20190412072055/https://www.ghacks.net/2015/04/22/create-publicly-available-web-page-archives-with-archive-is/) from the original on 12 April 2019. Retrieved 13 June 2015.
5. Brunelle, Justin F.; Kelly, Mat; Weigle, Michele C.; Nelson, Michael L. (25 January 2015). "The impact of JavaScript on archivability" (https://www.cs.odu.edu/~mweigle/papers/brunelle-ijdl16.pdf) (PDF). *International Journal on Digital Libraries*. **17** (2): 95–117. doi:10.1007/s00799-015-0140-8 (https://doi.org/10.1007%2Fs00799-015-0140-8). S2CID 8433375 (https://api.semanticscholar.org/CorpusID:8433375). Archived (https://web.archive.org/web/20190527064810/https://www.cs.odu.edu/~mweigle/papers/brunelle-ijdl16.pdf) (PDF) from the original on 27 May 2019.
6. Dascalescu, Dan (18 February 2013). "Web page archiving – Dan Dascalescu's Wiki (review)" (https://web.archive.org/web/20130922192354/http://wiki.dandascalescu.com/reviews/online_services/web_page_archiving). Wiki.dandascalescu.com. Archived from the original (https://wiki.dandascalescu.com/reviews/online_services/web_page_archiving) on 22 September 2013. Retrieved 3 October 2013.
7. Koebler, Jason (29 October 2014). "Dear GamerGate: Please Stop Stealing Our Shit" (https://motherboard.vice.com/read/dear-gamergate-please-stop-stealing-our-shit). *Motherboard*. Archived (https://archive.today/20190527064603/https://www.vice.com/en_us/article/ypw5mj/dear-gamergate-please-stop-stealing-our-shit) from the original on 27 May 2019. Retrieved 22 March 2017. "There is no way for a website to protect itself from having an Archive.today user mirror the site."
8. "Archive.today FAQ" (https://archive.today/faq). *archive.today*. Retrieved 15 February 2019.
9. "Home page of Archive.is in 2013" (https://web.archive.org/web/20130112221411/http://archive.is/). Archived from the original (http://archive.is/) on 12 January 2013.
10. "Archive.today blog" (https://blog.archive.today/post/657607767402659840/have-you-considered-allowing-small-mp4s-and-webms). Archived (https://web.archive.org/web/20210907153549/https://blog.archive.today/post/657607767402659840/have-you-considered-allowing-small-mp4s-and-webms) from the original on 7 September 2021.
11. *Archiving Websites with the Archive.is* (https://www.youtube.com/watch?v=LK_bp9_ZyQs), retrieved 27 January 2022
12. "Example snapshot history on archive.is" (https://archive.today/https://support.google.com/webmasters/answer/6062608?hl=en).

13. JavaScript-generated loading animation of Dailymotion video appearing in a frozen state (http s://archive.today/20200121182128/https://www.dailymotion.com/video/x3sexy8)

14. "Example: Page saved from Web Archive to Archive.is" (https://archive.vn/20130520191911/htt ps://es.wikipedia.org/wiki/Wikipedia) (in Spanish). Archived from the original (https://es.wikipedi a.org/wiki/Wikipedia) on 20 May 2013. Retrieved 23 October 2019.

15. For example, the string insite: https://en.wikipedia.org "World Cup" returns the "World+Cup"/ related snapshots (https://archive.today/search/?q=insite%3A+http%3Aen.wikipedia.org+)

16. "Some Frequently Asked Question" (https://blog.archive.is/post/41395737942/how-can-i-delete -an-archived-page) (blog). *archive.is*. 24 January 2013. Archived (https://web.archive.org/web/ 20130926093655/https://blog.archive.is/post/41395737942/how-can-i-delete-an-archived-pag e) from the original on 26 September 2013. Retrieved 12 November 2018.

17. "Example user request on the Archive.is blog" (https://blog.archive.today/post/6774275470641 56160/could-you-expand-17wus-and-links-under-same). *Archive.is blog*. Retrieved 7 April 2022.

18. "Example of dynamic list" (https://www.worldcat.org/search?q=au%3A%22thomas+aquinas%2 2&fq=&dblist=638&start=21&qt=page_number_link). *WorldCat.org*.

19. *Archiving Websites with the Archive.is* (https://www.youtube.com/watch?v=LK_bp9_ZyQs), retrieved 27 January 2022

20. "Just realized that I can search for keywords in the search bar for archive today, was this a recently added feature?" (https://blog.archive.today/post/673695282217762816/just-realized-th at-i-can-search-for-keywords-in). *Archive.is blog*. Retrieved 27 January 2022.

21. "Archive.is blog" (https://blog.archive.today/post/623883809154383872/the-download-zip-butto n-has-been-giving-a-not). 17 July 2020. Archived (https://web.archive.org/web/2020100312561 8/https://blog.archive.today/post/623883809154383872/the-download-zip-button-has-been-givi ng-a-not) from the original on 3 October 2020.

22. Nelson, Michael L. (9 July 2013). "Archive.is Supports Memento" (https://ws-dl.blogspot.nl/201 3/07/2013-07-09-archiveis-supports-memento.html). *Research and Teaching Updates*. Web Science and Digital Libraries Research Group at Old Dominion University. Archived (https://we b.archive.org/web/20130727194715/https://ws-dl.blogspot.de/2013/07/2013-07-09-archiveis-su pports-memento.html) from the original on 27 July 2013. Retrieved 17 September 2013.

23. "archive.is" (https://web.archive.org/web/20130915191950/https://mementoweb.org/depot/nativ e/archiveis/). *Memento Protocol Information*. Memento Development Group. Archived from the original (https://mementoweb.org/depot/native/archiveis/) on 15 September 2013. Retrieved 17 September 2013.

24. "Why did you change the URL back from archive-today to archive-is?" (https://blog.archive.is/p ost/118010496181/why-did-you-change-the-url-back-from-archive-today). *Archive.is Blog*. 3 May 2015. Archived (https://www.webcitation.org/6YwqBNeBD?url=http://blog.archive.is/post/1 18010496181/why-did-you-change-the-url-back-from-archive-today) from the original on 1 June 2015. Retrieved 6 January 2019.

25. @archiveis (4 January 2019). "Please do not use archive.IS mirror for linking, use others mirrors [.TODAY .FO .LI .VN .MD .PH]. .IS might stop working soon" (https://twitter.com/archive is/status/1081276424781287427) (Tweet). Archived (https://web.archive.org/web/2019010600 0101/https://twitter.com/archiveis/status/1081276424781287427) from the original on 6 January 2019 – via Twitter.

26. "ISPs in AU and NZ start censoring the internet without legal precedent" (https://www.privateint ernetaccess.com/blog/2019/03/isps-in-au-and-nz-start-censoring-the-internet-without-legal-pre cedent/). *Private Internet Access*. 19 March 2019. Retrieved 20 March 2019.

27. "New Zealand ISPs Say They're Blocking Sites That Fail To Remove Christchurch Shooting Video" (https://www.gizmodo.com.au/2019/03/new-zealand-isps-say-theyre-blocking-sites-that- fail-to-remove-christchurch-shooting-video/). *Gizmodo Australia*. 19 March 2019. Archived (http s://web.archive.org/web/20190518223849/https://www.gizmodo.com.au/2019/03/new-zealand-i sps-say-theyre-blocking-sites-that-fail-to-remove-christchurch-shooting-video/) from the original on 18 May 2019. Retrieved 20 March 2019.

28. "archive.is is 100% blocked in China" (https://en.greatfire.org/archive.is). *GreatFire Analyzer*. 12 August 2018. Archived (https://archive.today/20180812150852/https://en.greatfire.org/https/ archive.is) from the original on 12 August 2018.

29. "archive.li is 100% blocked in China" (https://en.greatfire.org/https/archive.li). *Great Fire Analyzer*. 12 August 2018. Archived (https://archive.today/20180812154815/https://en.greatfir e.org/https/archive.li) from the original on 12 August 2018.

30. "archive.fo is 100% blocked in China" (https://en.greatfire.org/https/archive.fo). *Great Fire Analyzer*. 12 August 2018. Archived (https://archive.today/20180812152220/https://en.greatfir e.org/https/archive.fo) from the original on 12 August 2018.

31. "archive.ph is 100% blocked in China" (https://en.greatfire.org/https/archive.ph). *en.greatfire.org*. Retrieved 7 April 2022.

32. Lapintie, Lassi (22 July 2015). "Suomalaisilta estettiin haktivistien suosimalla verkkosivulla käynti" (https://www.iltalehti.fi/digi/201507222007069_du.shtml) [Finns' access to website used by hacktivists blocked]. *Iltalehti* (in Finnish). Archived (https://web.archive.org/web/20190 527064017/https://www.iltalehti.fi/digi/a/201507222007069) from the original on 27 May 2019. Retrieved 4 March 2016.

33. Elistratov, Vladimir (29 January 2016). "Roskomnadzor zablokiroval servis archive.is, khranyashchiy kopii veb-saytov" Роскомнадзор заблокировал сервис archive.is, хранящий копии веб-сайтов (https://tjournal.ru/21966-roskomnadzor-zablokiroval-servis-archive-is-hrany ashchiy-kopii-veb-saytov). *TJournal* (in Russian). Archived (https://web.archive.org/web/20170 830055553/https://tjournal.ru/21966-roskomnadzor-zablokiroval-servis-archive-is-hranyashchiy -kopii-veb-saytov) from the original on 30 August 2017. Retrieved 30 January 2016.

34. Cushing, Tim (4 February 2016). "Russia Blocks Another Archive Site Because It Might Contain Old Pages About Drugs" (https://www.techdirt.com/articles/20160203/08365233504/ru ssia-blocks-another-archive-site-because-it-might-contain-old-pages-about-drugs.shtml). *Techdirt*. Archived (https://web.archive.org/web/20190323131754/https://www.techdirt.com/arti cles/20160203/08365233504/russia-blocks-another-archive-site-because-it-might-contain-old-p ages-about-drugs.shtml) from the original on 23 March 2019. Retrieved 26 February 2016.

35. "Archive.is – Error 1001" (https://community.cloudflare.com/t/archive-is-error-1001/18227). *Cloudflare Community*. 15 May 2018. Retrieved 2 December 2021.

36. @archiveis (16 July 2018). " "Having to do" is not so direct here. Absence of EDNS and massive mismatch (not only on AS/Country, but even on the continent level) of where DNS and related HTTP requests come from causes so many troubles so I consider EDNS-less requests from Cloudflare as invalid" (https://twitter.com/archiveis/status/1018691421182791680) (Tweet) – via Twitter.

37. "Comment by Matthew Prince on Hacker News" (https://archive.today/g9lCC). *Hacker News*. 4 May 2019. Archived from the original (https://news.ycombinator.com/item?id=19828702) on 13 May 2022. Retrieved 4 October 2021.

38. "Archive.today works again on 1.1.1.1 (and archive.{ph,is,li,vn,fo,md})" (https://community.clou dflare.com/t/archive-today-works-again-on-1-1-1-1-and-archive-ph-is-li-vn-fo-md/387566). *Cloudflare Community*. 22 May 2022. Retrieved 12 March 2023.

External links

- Official website (https://archive.today/) ✎
- Archive.is (https://archive-is.tumblr.com) on Tumblr
- archive.today (https://twitter.com/archiveis) on Twitter