

## Archiving the Intranet Resource List

**Note:** This list was compiled from responses to the following question, posted in the Business Archive Section's listserv (June, 2016): "Wondering if anyone out there has a good process/software/method for capturing material from your company's intranet?" Thank you to all who contributed in the creation of this list.

Tools used by larger institutions:

[Hanzo](#) - used by Coca Cola and other corporations for their public-facing site

- files are housed on Hanzo's servers (for compatibility reasons) but accessing them is easy (login/password)
- standard file formats of any given capture(s) can be requested
- User can select how many 'links deep' to go on a given site
- Works fairly well for capturing social media sites
  - Some structure is lost but the content is there and very easy to read
  - Does not capture comments

[Archive-It](#) - used primarily by Government and University Archives

Described as *Light-duty* Tools:

[HTTrack](#)

[WAR Create](#)

Tool for JavaScript-heavy, interactive content:

<http://webrecorder.io/>

- If you want the content saved on your own computer/servers, you can download the WARC files, generated by the tool
  - Note: this site was blocked by my firewall so I wasn't able to poke around on it.

Excellent Article exploring this topic:

[Leveraging Heritrix and the Wayback Machine on a Corporate Intranet: A Case Study on Improving Corporate Archives](#)

Resource for versions of Public Internet Sites:

[Internet Archive](#)

- [About the Internet Archive](#): "The Internet Archive is a 501(c)(3) non-profit that was founded to build an Internet library. Its purposes include offering permanent access for researchers, historians, scholars, people with disabilities, and the general public to historical collections that exist in digital format." (copied from the site, click link for full *About* description)

Date Created: July 20, 2016

Sources: Hive-mind of the Business Archives Section, SAA