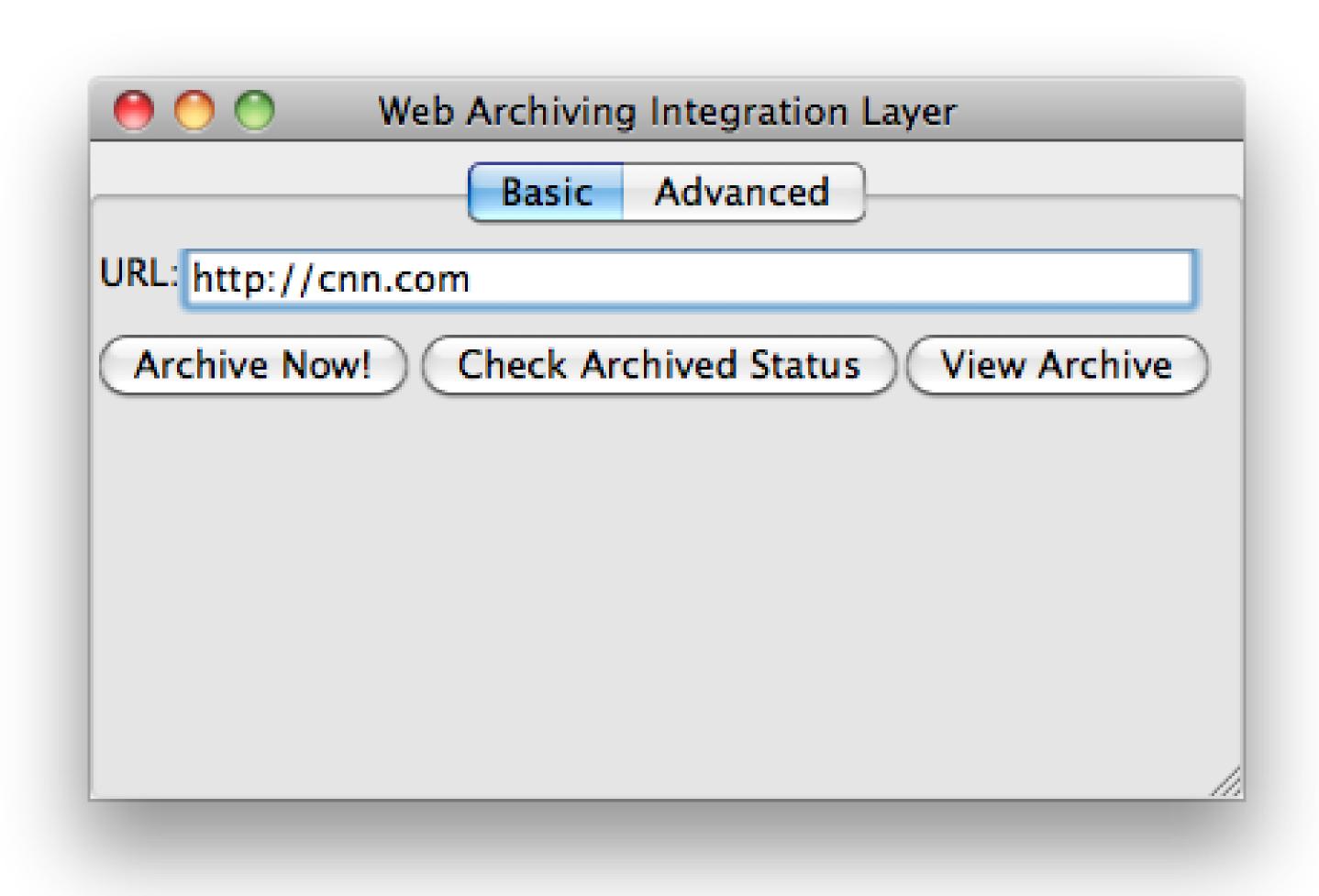


Making Enterprise-Level Archive Tools Accessible for Personal Web Archiving

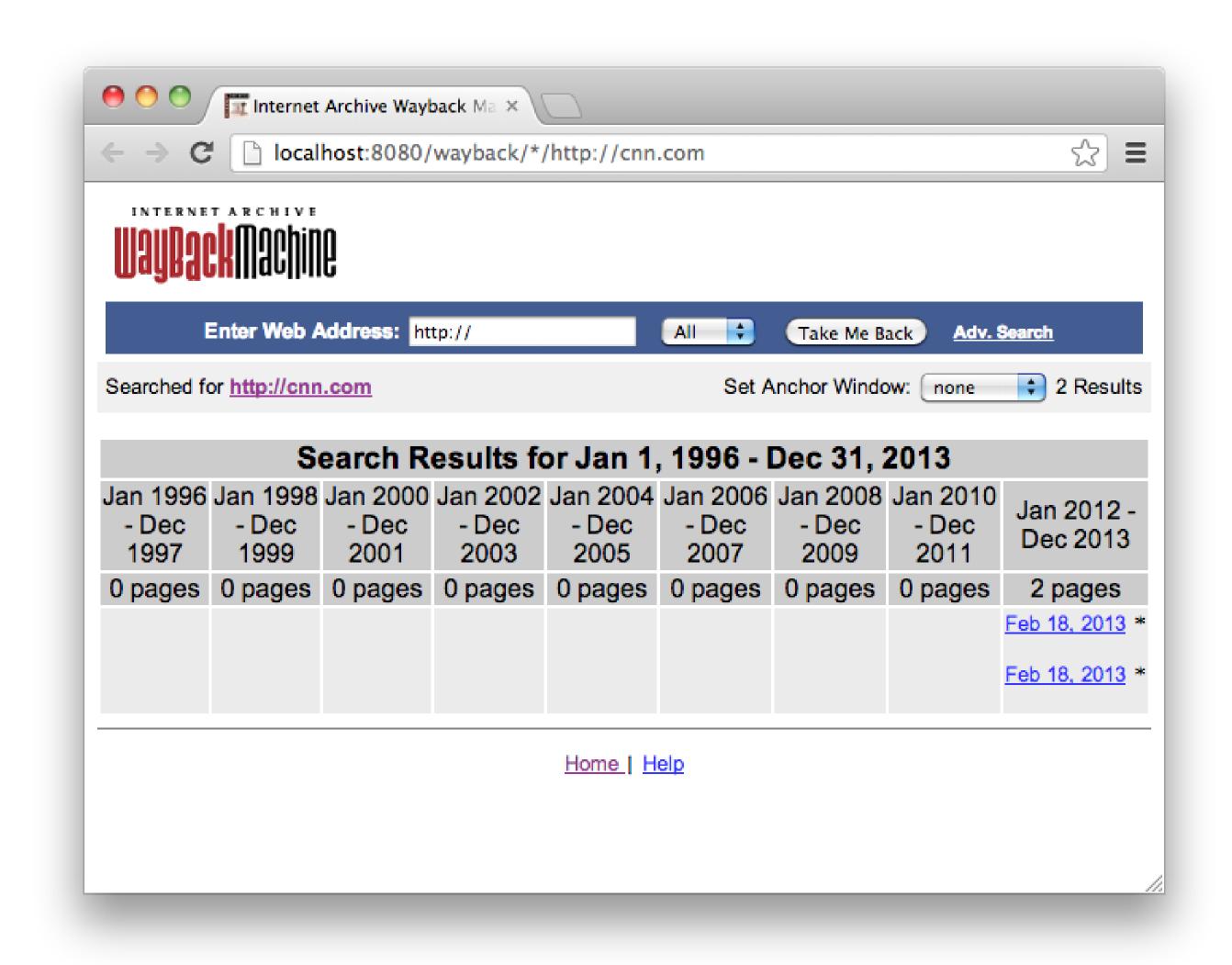
Mat Kelly, Michele C. Weigle, and Michael L. Nelson {mkelly, mweigle, mln}@cs.odu.edu
Department of Computer Science, Old Dominion University, Norfolk, Virginia USA



ONE-CLICK, USER INSTIGATED PRESERVATION



- "Archive Now!" button sets up crawl, initiates crawl and puts archive file in correct location to be indexed.
- Wayback consumption can be checked with "Check Archive Status" button.



• Once indexed, "View Archive" button shows all archives for URL in local Wayback.



• Selecting the date in local Wayback displays the preserved webpage.

FEATURES

- Collection of Archiving Tools
- Drag & Drop Installation And Removal
- All Tools Can Reside on a Single Machine
- Managed Through a Graphical User Interface (GUI)

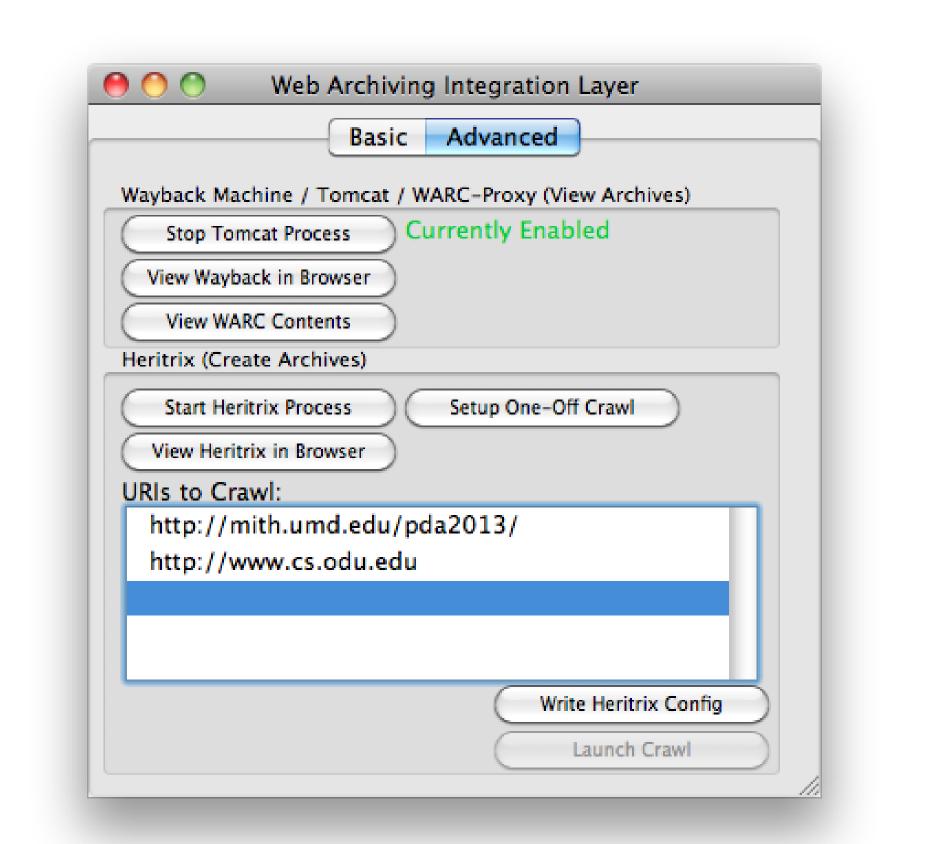
TOOLS INSTALLED LOCALLY

- CREATE ARCHIVES: Heritrix (Crawler)
- REPLAY ARCHIVES: Wayback Machine
- INSPECT ARCHIVES: WARC-Proxy
- More to Come!

Advanced Options/Features

- Specify Multiple URLs to be Included in the Crawl
- Setup Crawls and Allow for Customization Prior to Execution (e.g., crawl period)
- Start or Stop Services Not Currently Needed (e.g., initialize a long crawl but delay replay until later)

Interface for Tweaking



Support

- GENERATED ARCHIVES ARE SAFE
 Web ARChives (WARCs) reside on your hard
 drive, can be backed up for safe keeping like
 any other file
- CROSS PLATFORM
 Support for MacOS X, Windows and Linux
- WORKS WITH EXISTING WARCS

 Just drop in and local Wayback will index for replay
- COMPATIBLE WITH OTHER ARCHIVING TOOLS

Use the WARC-generating preservation tool of your choice (e.g., WARCreate, Wget) in lieu of Heritrix