

What kind of tcl tools would help me parse and use html info?

What kind of tcl tools would help me parse and use html info?

Source: <http://coding.derkeiler.com/Archive/Tcl/comp.lang.tcl/2006-03/msg01353.html>

- *From:* "Larry W. Virden" <lvirden@xxxxxxxxxx>
 - *Date:* 24 Mar 2006 04:10:56 -0800
-

I have a need to write a tool to do this:

fetch an html http URL

parse the html

Look through the A tags for some specific phrases

For each one found, check a file cache. If the URL associated with the tag is in the cache, see if it has been modified since it was placed into the cache. If not, continue.

If it has been modified, or if it doesn't exist in the cache, then

fetch the URL, place into the cache, and touch to make the cache copy have the date and time from the web site.

For one of the specific phrases, instead of caching the file, treat it as the next html to parse and search.

When one specific term is no longer found, application is finished.

The only other possible thing for the algorithm above is that one of the URLs is the URL of a CGI with values. The other URLs are just static HTML pages.

What are some examples using some of the Tcl tools for parsing that fetched file and searching the A tags for phrases?

.