Editing Netscape Bookmark Files

Netscape bookmark files contain extraneous information which can be easily edited out of the HTML file before importing it into HTML.edit. This takes the form of titles, comments, H3 headings, list item tags, and ADD_DATE and LAST_VISIT attributes on anchors. The trick is learning Grep.

Blech. Barf. Yuk. (to quote Dan Connolly) Maybe BBEdit and I can help. While the format of a Netscape bookmark file may change (this description fits version 1.1N), the following explanation will get you started. When performing a replace, turn on Grep where indicated and use the recommended "Search for:" and "Replace" with: strings.

1. Delete the first two lines: `<!DOCTYPE>` and an HTML comment.

2. Remove all `<P>`, `<HR>`, `<DL>`, `</DL>` and `<DT>` tags (Match Case and Grep turned off).

3. Any existing subheadings are instantiated as <H3> elements. Remove the tags from around the text and prefix each on a line with `(>I` to disable and make it italic (Grep on):

    Search for: `<[hH]3 ([^>]*)>`
    Replace with: `(>I`

Remove all instances of `</H3>`.

4. Remove all ADD_DATE and LAST_VISIT attributes from the remaining anchors:

    Search for: `ADD_DATE="[-0-9]*"`
    Search for: `LAST_VISIT="[-0-9]*"`

Prefix each Search for: string with a space, and replace in each case with nothing (Grep on).

5. Remove all preceding and trailing spaces on each line (Grep on):

    Search for: `\r[ ]*`
    Replace with: `\r`

    Search for: `[ ]*\r`
    Replace with: `\r`

Then just trim out the empty lines and you're done.

While this process took some time to figure out initially, it now takes me about five minutes to translate a Hotlist file, and I thought I'd pass this tidbit along. Since BBEdit is AppleScriptable, I suppose someone could script this. (Anyone volunteer?)

Back to Hotlist Maintenance, Libraries Preferences, or return to HTML.edit Features.