InstallingWeb2

NOTE: Web2 module is no longer part of Nutch

So these instructions do no longer apply.

---

chris sleeman wrote:
> Hi,
>
> Can anyone tell me how to use the spell-check query plugin available in the
> contrib \ web2 dir (and even the rest of the plugins too)? Is it similar to
> enabling the nutch-plugins?

Following these steps should get you there:

   1. compile nutch (in top level dir do "ant")

2. crawl your data (see tutorial)

3. edit your conf/nutch-site.xml so it contains plugin "web-query-propose-spellcheck" and "webui-extensionpoints"

4. edit conf/nutch-site.xml so it contains proper dir for plugins as the plugins are not packaged inside .war (something like
<property>
<name>plugin.folders</name>
<value> <path to plugins dir> </value>
</property>
)

5. compile web2 plugins (in contrib/web2 do ant compile-plugins)

6. edit search.jsp contains line "<tiles:insert definition="propose" ignore="true"/>" just before the second c:choose.

7. create web2 app (in contrib/web2 do ant war)

8. build your spell check index (bin/nutch plugin web-query-propose-spellcheck org.apache.nutch.spell.NGramSpeller -i
<indexdir> -f content -o spelling

9. deploy webapp to tomcat

10. start tomcat (from the dir you have your crawl data and ngram index generated in #7)

11. search for something that is spelled incorrectly

> Also how do we build the spelling index ? Are these plugins still "WIP" ? I

see #8 above, the whole web is MWSN (More Work Still Needed🙂)

> haven't been able to find any docs on these.

That's because there currently is not any other documentation but the readme in http://svn.apache.org/viewvc/lucene/nutch/trunk/contrib/web2/README.txt?view=markup

I should probably put some documentation to wiki to gain more attraction

fyi - I just committed a small fix to bug that might prevent spell checking proposer from working. So if you have problems check out the trunk or a nightly build tomorrow.

–
Sami Siren