**Steps for Protocol-Interactive Selenium plugin**

1. Added protocol-interactive selenium in plugin.includes property of nutch-site.xml.

<property>

<name>plugin.includes</name>

<value>protocol-interactiveselenium|protocol-http|urlfilter-regex|parse-(tika|html)|index-(basic|anchor)|indexer-solr|scoring-similarity|urlnormalizer-(pass|regex|basic)</value>

</property>

1. Commented out (.\*[?\*!@=].\*) in /runtime/local/conf/regex-urlfilter.txt and automaton-urlfilter.txt to allow URLs containing certain characters as probable queries.
2. Extended and developed a plugin for protocol-interactive-selenium to resolve issues encountered in 1st crawl like data behind forms, Ajax based interaction, and JavaScript based pagination etc. using selenium WebDriver.
3. Rebuild the nutch (ant runtime) with custom plugin.
4. Added custom plugin in interactiveselenium.handlers property in nutch-site.xml.

<property>

<name>interactiveselenium.handlers</name>

<value>CrawlHandler,DefaultHandler</value>

<description></description>

</property>

1. Re-run the crawl.