Clustering web search results

Published April 21, 2007

Citation

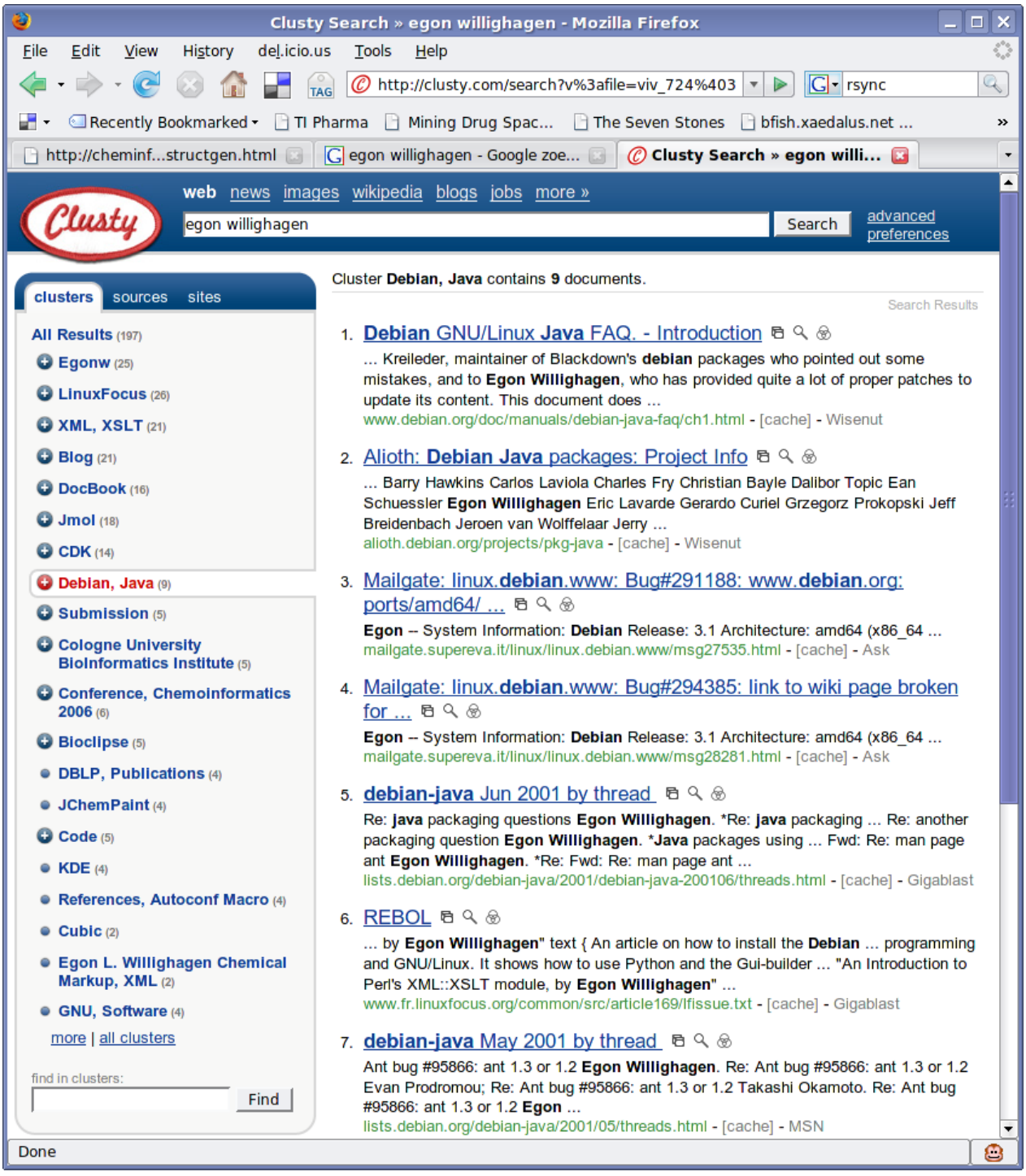
V2lsbGlnaGFnZW4sIEUuICgyMDA3LCBBcHJpbCAyMSkuIENsdXN0ZXJpbmcgd2ViIHNlYXJjaCBy
ZXN1bHRzLiA8aT5DaGVtLWJsYS1pY3M8L2k+LiBodHRwczovL2RvaS5vcmcvMTAuNTkzNTAveXRw
c3MtZnc2NjY=

Keywords
Google, Cheminf

Abstract
The Dutch Intermediair magazine of this week had a letter sent by a reader introducing Clusty, a web search engine that clusters the results.

chem-bla-ics

The Dutch Intermediair magazine of this week had a letter sent by a reader introducing Clusty, a web search engine that clusters the results. It does a pretty good job for 'egon willighagen':



It seems to use other engine to do the searching and focus on the clustering. Source engine exclude Google, and include Gigablast, MSN and Wikipedia.

For *chemoinformatics* it comes up with the following top 10 clusters: 'Drug Discovery', 'Structure', 'Cheminformatics', 'Research', 'Books', 'Conference, German', 'Textbook, Gasteiger', 'Laboratory', 'Handbook of Chemoinformatics', and 'School'. Quite acceptable and useful clustering.

chem-bla-ics

This might be the next step in googling. Rich, it also might solve your problem : searching for ‘ruby chemoinformatics’ does **not** give a ‘Depth First’ or ‘Rich Apodaca’ cluster :)