CATS: We will only use one method for cats since it will not make sense to first check for sets in the individual article when only working with one article.

Query	Method regex matches	Method grep matches	Time regex	Time grep
['kitten', (15, 85), 'cat', (0, 100), 'sire', (0, 200), 'oxford']	1	1	0.022s	0.018s
['english', (0,200), 'cat']	12	12	0.027s	0.02s
['china', (30, 150), 'washington']	2	1	0.02s	0.013s
['cat', (0, 100), 'anatomy']	4	4	0.024s	0.017s
['cat', (0, 10), 'are', (0, 10), 'to']	9	9	0.015s	0.012s

A articles: We were not able to select the A articles in the big wiki dump file.

Query	Method 1 matches	Method 2 matches	Time 1	Time 2

All articles:

Query	Method regex matches	Method grep matches	Time regex	Time grep
['elephants', (0, 20), 'are', (0, 20), 'to']	238	238	6:31m	5:49m
['technical', (0, 20), 'university', (0, 20), 'denmark']	15	15	5:54m	5:27m
['testing', (0, 20), 'with', (0, 20), 'a', (0, 30), 'lot', (0, 4), 'of', (0, 5), 'words']	0	0	8:44m	7:22m
['stress', (0, 250), 'test']	1640	1640	7:52m	6:14m
['object', (10, 200), 'application', (0, 100), 'python', (10, 200), 'system', (0, 100), 'computer', (0, 10), 'science', (0, 150), 'linux', (0, 200), 'ruby']	0	0	11:37m	9:11m