

crawl web

# Search Engine

crawler

Monitor  
url list

unknown

URL 1	Simhash Pattern
URL 2	Simhash Pattern
URL 3	Simhash Pattern
URL 4	Simhash Pattern
...	...

ok

bad

Simhash Blacklist

update list

submit

www.example.com

RAPPOR

url  
simhash

Need to inspect?  
ie. from search engine?

unknown

compute  
simhash

check

Blacklist

Whitelist

ok

bad

Cloaking Detction Extension