

Project 3 M1: Initial Index constructions and Retrieval

Milestone #1 Deliverable

Team Number: 6

	Group Member #1	Group Member #2	Group Member #3
Name	Keyu Zhang	Chak wah Lo	Emanuel Lopez
UCI NET ID	keyuz4	cwlo1	emanuel1

#1) Index Statistic:

Fill out the table below with the appropriate information pertaining to your groups index.

Number of documents	_____37497_____ documents
Number of unique words	_____510644_____ unique words
Total size in kb	_____123764_____ KB

#2) Sample Query Results:

Number of URLs retrieved for each of the queries above and listing the first 20 URLs for each query

Term	Num ber of Links	URLs
Informa tics	6095	wics.ics.uci.edu/category/news/info-session?afg12_page_id=2&afg44_page_id=4&afg34_page_id=4&afg15_page_id=2&afg42_page_id=2&afg22_page_id=2 www.ics.uci.edu/faculty/profiles/view_faculty.php/grad/degrees/index/faculty/centers/ugrad/sao/centers/ugrad/index/involved/about/ugrad/policies/Withdrawal_Readmission/faculty/centers/about/search/search_graduate_all.php/grad/policies/index/about/about_contact.p hp/about/about_deanmsg.php/grad/funding/index www.ics.uci.edu/alumni/mentor/hall_of_fame/hall_of_fame/stayconnected/hall_of_fame/stayconnected/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/hall_of_fame/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/stayco nnected/hall_of_fame/inductees.aspx.php www.ics.uci.edu/alumni/mentor/hall_of_fame/hall_of_fame/stayconnected/hall_of_fame/stayconnected/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/hall

		<p>_of_fame/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/stayco nnected/hall_of_fame/stayconnected/stayconnected/stayconnected/stayconnected/stayc onected/stayconnected/hall_of_fame/stayconnected/stayconnected/hall_of_fame/hall_of_ _fame/stayconnected/stayconnected/stayconnected/stayconnected/hall_of_fame/staycon nected/hall_of_fame/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_f ame/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/stayconnect ed/stayconnected/stayconnected/stayconnected/index.php</p> <p>wics.ics.uci.edu/category/news/info-session?afg12_page_id=2page%2F2%2Fpage/2/&afg 42_page_id=2&afg44_page_id=4&afg43_page_id=3&afg34_page_id=4</p> <p>www.ics.uci.edu/ugrad/degreess/Doube_Majors.php</p> <p>www.ics.uci.edu/about/visit/./bren/index.php/bren_advance.php/./brenhall/index.php/bre n_vision.php/././ugrad/index.php/bren_vision.php/bren_press.php/./about_mission.php/ ./about_safety.php/index.php/bren_advance.php/./search/index.php/bren_press.php/bren _vision.php/./brenhall/index.php/./visit/index.php/././ugrad/index.php/./about_safety.php /bren_vision.php/./visit/index.php/././ugrad/index.php/./about_contact.php/bren_about.p hp/./brenhall/index.php/bren_vision.php/bren_press.php/index.php/bren_press.php/./sea rch/index.php/./search/index.php/ICS/ics/about/about_facts/././ugrad/index.php/././ugra d/index.php/././ugrad/index.php/ICS/ics/about/about_facts/bren_press.php/././grad/index .php/./about_deanmsg.php/./about_deanmsg.php/./about_mission.php/./visit/index.php/ ./about_contact.php/bren_about.php</p> <p>www.ics.uci.edu/faculty/profiles/view_faculty.php/faculty/centers/faculty/about/about_con tact.php/about/about_contact.php/ugrad/QA_Petitions/community</p> <p>www.ics.uci.edu/facebook/login/browse/fanned_pages/events/252272701892321/search/st r/%25E9%259D%259E%25E8%2590%25A5%25E5%2588%25A9%25E7%25BB%2584%25E7 %25BB%2587/keywords_pages/UCIBrenICS/photos/a.10150104180706909.283411.1062914 86908/10154408253796909?type=3</p> <p>www.ics.uci.edu/alumni/mentor/stayconnected/stayconnected/stayconnected/hall_of_fam e/stayconnected/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/ stayconnected/stayconnected/index.php</p> <p>sdcl.ics.uci.edu/2015/06/end-of-quarter-lunch-2</p> <p>www.ics.uci.edu/alumni/hall_of_fame/hall_of_fame/stayconnected/stayconnected/hall_of_ fame/hall_of_fame/hall_of_fame/stayconnected/stayconnected/hall_of_fame/hall_of_fame/ stayconnected/stayconnected/hall_of_fame/hall_of_fame/stayconnected/stayconnected/s tayconnected/index.php</p> <p>www.ics.uci.edu/alumni/stayconnected/stayconnected/hall_of_fame/hall_of_fame/staycon nected/stayconnected/hall_of_fame/hall_of_fame/hall_of_fame/stayconnected/hall_of_fa me/hall_of_fame/hall_of_fame/stayconnected/index.php</p> <p>www.ics.uci.edu/~kay/courses/i42/hw</p> <p>www.ics.uci.edu/alumni/mentor/hall_of_fame/hall_of_fame/stayconnected/hall_of_fame/st ayconnected/stayconnected/stayconnected/hall_of_fame/stayconnected/hall_of_fame/sta yconnected/stayconnected/stayconnected/index.php</p> <p>www.ics.uci.edu/facebook/pages/create/dialog/UCIBrenICS/videos/10154385262331909/ev ents/1849114518693808/browse/fanned_pages?id=106291486908&showauxiliary=1&fanori gin=timeline_like_chaining&ref=page_internal</p> <p>sdcl.ics.uci.edu/2013/12/new-workshop-on-crowdsourcing-in-software-engineering</p> <p>www.ics.uci.edu/about/visit/./bren/index.php/bren_advance.php/./brenhall/index.php/bre n_vision.php/././ugrad/index.php/bren_vision.php/bren_press.php/./about_mission.php/ ./about_safety.php/index.php/bren_advance.php/./search/index.php/bren_press.php/bren _vision.php/./brenhall/index.php/./visit/index.php/././ugrad/index.php/./about_safety.php /bren_vision.php/./visit/index.php/././ugrad/index.php/./about_contact.php/bren_about.p hp/./brenhall/index.php/bren_vision.php/bren_press.php/index.php/bren_press.php/./sea rch/index.php/./search/index.php/ICS/ics/about/about_facts/././ugrad/index.php/././ugra d/index.php/././ugrad/index.php/ICS/ics/about/about_facts/bren_press.php/././grad/index .php/./about_deanmsg.php/./about_deanmsg.php/./about_mission.php/./visit/index.php/ ./././ugrad/index.php/bren_press.php/index.php/bren_advance.php/index.php/bren_about. php/./about_mission.php/./about_deanmsg.php/././ugrad/index.php/index.php/./search/i ndex.php/bren_about.php/./search/index.php/./visit/index.php/bren_about.php</p> <p>www.ics.uci.edu/faculty/profiles/view_faculty.php/faculty/centers/faculty/about/about_con tact.php/about/search/index.php/grad/policies/index/involvement/project_class</p> <p>www.ics.uci.edu/faculty/area/area_compdesign.php</p>
Mondeg o	50	<p>mondego.ics.uci.edu/projects/clonedetection/tool/source?C=S;O=A</p> <p>www.ics.uci.edu/community/news/notes/notes_2013.php</p> <p>mondego.ics.uci.edu/projects/clonedetection/tool/demo</p> <p>mondego.ics.uci.edu/projects/clonedetection/tool/latest/features?C=D;O=A</p> <p>mondego.ics.uci.edu/datasets/wikipedia-events/files?C=M;O=A</p> <p>mondego.ics.uci.edu/projects/clonedetection/tool/demo?C=S;O=A</p> <p>mondego.ics.uci.edu/projects/clonedetection/tool/latest/plugins</p> <p>mondego.ics.uci.edu/datasets/wikipedia-events/files?C=M;O=D</p> <p>mondego.ics.uci.edu</p> <p>mondego.ics.uci.edu/datasets?C=N;O=D</p> <p>www.ics.uci.edu/~lopes</p>

[illegible]