

# Project 3 M1: Initial Index constructions and Retrieval

## Milestone #1 Deliverable

**Team Number: 39/44**

Course Groups

[Project 2 Crawler 10](#)

[Project 3 Search System 39](#)

[Project 3 Search System 44](#)

	Group Member #1	Group Member #2	Group Member #3
Name	Zion Mateo	Matthew Chang	Steven Tu
UCI NET ID	mateoz	mcchang2	stu7

### #1) Index Statistic:

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

Number of unique document ids in index	37434 documents
Number of unique words	6886149 unique words
Total size in kb	25231326 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	Number of Links	URLs
------	-----------------	------

Informatics	5965	<p>1. URL:  <a href="http://www.ics.uci.edu/faculty">www.ics.uci.edu/faculty</a></p> <p>2. URL:  <a href="http://www.ics.uci.edu/faculty/index.php">www.ics.uci.edu/faculty/index.php</a></p> <p>3. URL:  <a href="http://www.ics.uci.edu/faculty/index.php?department=Informatics">www.ics.uci.edu/faculty/index.php?department=Informatics</a></p> <p>4. URL:  <a href="http://www.ics.uci.edu/~kay/courses/previous.html">www.ics.uci.edu/~kay/courses/previous.html</a></p> <p>5. URL:  <a href="http://www.ics.uci.edu/ugrad/degrees/degree_in4matx.php">www.ics.uci.edu/ugrad/degrees/degree_in4matx.php</a></p> <p>6. URL:  <a href="http://www.ics.uci.edu/~andre/teaching.html">www.ics.uci.edu/~andre/teaching.html</a></p> <p>7. URL:  <a href="http://www.ics.uci.edu/community/news/notes/notes_2008.php">www.ics.uci.edu/community/news/notes/notes_2008.php</a></p> <p>8. URL:  <a href="http://www.ics.uci.edu/about/annualreport/2006-07/events.php">www.ics.uci.edu/about/annualreport/2006-07/events.php</a></p> <p>9. URL:  <a href="http://www.ics.uci.edu/community/news/spotlight/spotlight_purpura.php">www.ics.uci.edu/community/news/spotlight/spotlight_purpura.php</a></p> <p>10. URL:</p>
-------------	------	--

		<p><a href="http://www.ics.uci.edu/community/news/notes/notes_2014.php">www.ics.uci.edu/community/news/notes/notes_2014.php</a></p> <p>11. URL: <a href="http://www.ics.uci.edu/ugrad/qa">www.ics.uci.edu/ugrad/qa</a></p> <p>12. URL: <a href="http://www.ics.uci.edu/community/news/notes/notes_2010.php">www.ics.uci.edu/community/news/notes/notes_2010.php</a></p> <p>13. URL: <a href="http://www.ics.uci.edu/prospective/en/degrees/informatics">www.ics.uci.edu/prospective/en/degrees/informatics</a></p> <p>14. URL: <a href="http://www.ics.uci.edu/~kay/courses/131/su12.html">www.ics.uci.edu/~kay/courses/131/su12.html</a></p> <p>15. URL: <a href="http://www.ics.uci.edu/~kay/courses/131/su13.html">www.ics.uci.edu/~kay/courses/131/su13.html</a></p> <p>16. URL: <a href="http://www.ics.uci.edu/~kay/courses/131/su14.html">www.ics.uci.edu/~kay/courses/131/su14.html</a></p> <p>17. URL: <a href="http://www.ics.uci.edu/~kay/courses/131">www.ics.uci.edu/~kay/courses/131</a></p> <p>18. URL: <a href="http://www.ics.uci.edu/community/news/spotlight/spotlight_hornick.php">www.ics.uci.edu/community/news/spotlight/spotlight_hornick.php</a></p> <p>19. URL: <a href="http://www.ics.uci.edu/~kay/courses/131/su11.html">www.ics.uci.edu/~kay/courses/131/su11.html</a></p>
--	--	--

		<b>20. URL:</b> <a href="http://www.ics.uci.edu/~andre/informatics209sf2006.html">www.ics.uci.edu/~andre/informatics209sf2006.html</a>
<b>Mondego</b>	<b>6</b>	<b>1. URL:</b> <a href="http://mondego.ics.uci.edu">mondego.ics.uci.edu</a> <b>2. URL:</b> <a href="http://www.ics.uci.edu/community/news/notes/notes_2013.php">www.ics.uci.edu/community/news/notes/notes_2013.php</a> <b>3. URL:</b> <a href="http://www.ics.uci.edu/~lopes">www.ics.uci.edu/~lopes</a> <b>4. URL:</b> <a href="http://www.ics.uci.edu/community/news/notes/notes_2014.php">www.ics.uci.edu/community/news/notes/notes_2014.php</a> <b>5. URL:</b> <a href="http://www.ics.uci.edu/~aftabh">www.ics.uci.edu/~aftabh</a> <b>6. URL:</b> <a href="http://sourcerer.ics.uci.edu">sourcerer.ics.uci.edu</a>
<b>Irvine</b>	<b>11535</b>	<b>1. URL:</b> <a href="http://www.ics.uci.edu/~eppstein/bibs/eppstein.html">www.ics.uci.edu/~eppstein/bibs/eppstein.html</a> <b>2. URL:</b> <a href="http://www.ics.uci.edu/~eppstein/bibs/eppstein.bib">www.ics.uci.edu/~eppstein/bibs/eppstein.bib</a> <b>3. URL:</b> <a href="http://cml.ics.uci.edu/category/aiml">cml.ics.uci.edu/category/aiml</a> <b>4. URL:</b> <a href="http://www.ics.uci.edu/facebook">www.ics.uci.edu/facebook</a>

		<p><b>/login/browse/fanned_pages/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages?ref=page_about_category 5.</b></p> <p><b>URL:</b></p> <p><b>www.ics.uci.edu/facebook/pages/create/dialog/UCIBrenICS/videos/10154385262331909/search/112204368792315/places-in/199405806739848/places/intersect/events/711137452397667?acontext=%7B%22source%22%3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22</b></p>
--	--	---

		<p>%2C%22mechanism%22 %3A%22feed_story%22% 2C%22extra_data%22%3 A%5B%5D%7D%5D&amp;has_ source=1&amp;hc_ref=PAGES _TIMELINE</p> <p>6. URL: www.ics.uci.edu/facebook /browse/likes/login/UCIBr enICS/photos/a.10150104 180706909.283411.106291 486908/101544027065419 09?type=3</p> <p>7. URL: www.ics.uci.edu/facebook /login/browse/fanned_pag es/search/str/%25E6%259 5%2599%25E8%2582%25 B2/keywords_pages/UCIB renICS/photos/a.10150104 180706909.283411.106291 486908/101543512033069 09?type=3</p> <p>8. URL: www.ics.uci.edu/facebook /events/340846149648374/ UCIReunion2017/photos/o .106291486908/262134790 874470/browse/fanned_pa ges/login/search/str/%25E 6%2595%2599%25E8%25 82%25B2/keywords_page</p>
--	--	--

		<p>s?ref=page_about_category</p> <p>9. URL:  <a href="http://www.ics.uci.edu/facebook/pages/create/dialog/UCIBrenlCS/photos/a.10150104180706909.283411.106291486908/10154351203306909?type=3">www.ics.uci.edu/facebook/pages/create/dialog/UCIBrenlCS/photos/a.10150104180706909.283411.106291486908/10154351203306909?type=3</a></p> <p>10. URL:  <a href="http://www.ics.uci.edu/facebook/events/340846149648374?context=%7B%22ref%22%3A%2298%22%2C%22action_history%22%3A%22null%22%7D">www.ics.uci.edu/facebook/events/340846149648374?context=%7B%22ref%22%3A%2298%22%2C%22action_history%22%3A%22null%22%7D</a></p> <p>11. URL:  <a href="http://www.ics.uci.edu/facebook/events/340846149648374/UCIReunion2017/photos/o.106291486908/262134790874470/browse/fanned_pages/UCIBrenlCS/videos/10154385262331909">www.ics.uci.edu/facebook/events/340846149648374/UCIReunion2017/photos/o.106291486908/262134790874470/browse/fanned_pages/UCIBrenlCS/videos/10154385262331909</a></p> <p>12. URL:  <a href="http://www.ics.uci.edu/facebook/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages/events/246966085750971/UCIBrenlCS/photos/a.10150104180706909.283411.1062914">www.ics.uci.edu/facebook/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages/events/246966085750971/UCIBrenlCS/photos/a.10150104180706909.283411.1062914</a></p>
--	--	---

		<p>86908/10154351203306909/browse/fanned_pages?id=106291486908&amp;showauxiliary=1&amp;fanorigin=timeline_like_chaining&amp;ref=page_internal</p> <p>13. URL: www.ics.uci.edu/facebook/login/browse/fanned_pages/events/252272701892321/search/str/%25E9%259D%259E%25E8%2590%25A5%25E5%2588%25A9%25E7%25BB%2584%25E7%25BB%2587/keywords_pages/UCIBrenICS/photos/a.10150104180706909.283411.106291486908/10154408253796909?type=3</p> <p>14. URL: www.ics.uci.edu/facebook/login/browse/fanned_pages/events/252272701892321?accontext=%22%3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism</p>
--	--	--



		<p><b>%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22%2C%22mechanism%22%3A%22feed_story%22%2C%22extra_data%22%3A%5B%5D%7D%5D&amp;has_source=1&amp;hc_ref=PAGES_TIMELINE</b></p> <p><b>15. URL:</b> <b>www.ics.uci.edu/facebook/UCIBrenICS/photos/a.10150104180706909.283411.106291486908/10154402706541909?type=3</b></p> <p><b>16. URL:</b> <b>www.ics.uci.edu/facebook/events/340846149648374/UCIReunion2017/photos/o.106291486908/262134790874470/browse/fanned_pages/UCIBrenICS/about?__xt__=33.%7B%22logging_data%22%3A%7B%22page_id%22%3A%221062914</b></p>
--	--	---

		<p>86908%22%2C%22event_type%22%3A%22clicked_view_page_about%22%2C%22impression_info%22%3A%22%22%2C%22surface%22%3A%22www_pages_highlights%22%2C%22interacted_story_type%22%3A%22291992504486569%22%2C%22session_id%22%3A%223baeaff90e3d466c803e51114282bdbc%22%7D%7D</p> <p>17. URL:  <a href="http://www.ics.uci.edu/facebook/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages?ref=page_about_category">www.ics.uci.edu/facebook/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages?ref=page_about_category</a></p> <p>18. URL:  <a href="http://www.ics.uci.edu/facebook/browse/likes/UCIBrenICS/photos/pb.106291486908.-2207520000.1487108481./10154420875021909/search/str/%25E9%259D%259E%25E8%2590%25A5%25E5%2588%25A9%25E7%25BB%2584%25E7%25BB%2587/keywords_pages/events/252272701892321?acontext=%7B%22source%">www.ics.uci.edu/facebook/browse/likes/UCIBrenICS/photos/pb.106291486908.-2207520000.1487108481./10154420875021909/search/str/%25E9%259D%259E%25E8%2590%25A5%25E5%2588%25A9%25E7%25BB%2584%25E7%25BB%2587/keywords_pages/events/252272701892321?acontext=%7B%22source%</a></p>
--	--	--

		<p>22%3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22%2C%22mechanism%22%3A%22feed_story%22%2C%22extra_data%22%3A%5B%5D%7D%5D&amp;has_source=1&amp;hc_ref=PAGES_TIMELINE</p> <p>19. 35/65, URL: <a href="http://www.ics.uci.edu/facebook/browse/likes/UCIBrenICS/photos/pb.106291486908.-2207520000.1487108481./10154420875021909/events/711137452397667?acontext=%5B%22source%22%22%3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22%2C%22mechanism%22%3A%22feed_story%22%2C%22extra_data%22%3A%5B%5D%7D%5D&amp;has_source=1&amp;hc_ref=PAGES_TIMELINE">www.ics.uci.edu/facebook/browse/likes/UCIBrenICS/photos/pb.106291486908.-2207520000.1487108481./10154420875021909/events/711137452397667?acontext=%5B%22source%22%22%3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22%2C%22mechanism%22%3A%22feed_story%22%2C%22extra_data%22%3A%5B%5D%7D%5D&amp;has_source=1&amp;hc_ref=PAGES_TIMELINE</a></p>
--	--	--

		<p>3A3%2C%22source_newsfeed_story_type%22%3A%22regular%22%2C%22action_history%22%3A%22%5B%7B%5C%22surface%5C%22%3A%5C%22newsfeed%5C%22%2C%5C%22mechanism%5C%22%3A%5C%22feed_story%5C%22%2C%5C%22extra_data%5C%22%3A%5B%5D%7D%5D%22%2C%22has_source%22%3Atrue%7D&amp;source=3&amp;source_newsfeed_story_type=regular&amp;action_history=%5B%7B%22surface%22%3A%22newsfeed%22%2C%22mechanism%22%3A%22feed_story%22%2C%22extra_data%22%3A%5B%5D%7D%5D&amp;has_source=1&amp;hc_ref=PAGES_TIMELINE</p> <p>20. URL: <a href="http://www.ics.uci.edu/facebook/browse/likes/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages/UCIBrenlCS/photos/a.10150104180706909.283411.106291486908/10154">www.ics.uci.edu/facebook/browse/likes/search/str/%25E6%2595%2599%25E8%2582%25B2/keywords_pages/UCIBrenlCS/photos/a.10150104180706909.283411.106291486908/10154</a></p>
--	--	--

		<b>351203306909?type=3</b>
--	--	----------------------------