

Errata sheet

Document type: Master's Thesis in Computer Science
Title: Incremental Information Retrieval
Subtitle: Finding new information by registering and ignoring already seen search results
Author: Erlend Johannessen
Year: June 2017
Department: Faculty of Science and Technology, Department of Computer Science
University: UiT, the Arctic University of Norway
Errata date: June 11, 2019

Data

Page	Problem	Correction
132	In appendix A, in table A.3, the three columns of averages "Avg", "Avg" and "Avg d2" all have wrong values.	Divisor was 71. The correct divisor should be 54, the selected number of days in the period. Correcting the divisor resulted in somewhat larger averages for these three columns. See the updated table A.3.

Language

Page	Line	Needs correction	Corrected by
Abstract	3	result quality <u>off</u> this kind	result quality <u>of</u> this kind
1	4	<u>organization</u> of	<u>organisation</u> of
12	18	highly <u>optimized</u>	highly <u>optimised</u>
13	3	or <u>personalized</u> results	or <u>personalised</u> results
14	2	and <u>trustable</u> videos	and <u>trustworthy</u> videos
15	7	Incremental Information Retrieval (<u>IIR</u>)	Incremental Information Retrieval (<u>IIR</u>)
17, 114, 119	13, 4, 8	<u>etcetera</u>	<u>et cetera</u>
47	20	<u>parametrised</u> executable	<u>parameterised</u> executable
89	4	happen to the <u>folowing</u>	happen to the <u>following</u>
102	15	<u>useable</u>	<u>usable</u>
104	3	which was <u>an</u> good fit	which was <u>a</u> good fit
109	26	<u>An</u> an example, try	<u>As</u> an example, try
115	16	<u>behavior</u>	<u>behaviour</u>
117	8	maintenance, <u>bugfixes</u>	maintenance, <u>bug fixes</u>

ID	Query	Exact	Results	Avg	Min	Max	New	New %	Avg	Min	Max	New d2	New d2 %	Avg d2
1	Messerschmitt KR200 restoration	—	40 194	744.3	612	846	1596	4.0	29.6	0	509	534	1.3	10.1
2	Messerschmitt KR200 restoration	Yes	20 451	378.7	285	478	42	0.2	0.8	0	18	9	0.0	0.2
3	Web search API thesis	—	50 919	942.9	646	1000	3111	6.1	57.6	0	770	1087	2.1	20.5
4	Web search API thesis	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
5	Web search thesis	—	52 143	965.6	809	998	3164	6.1	58.6	0	871	1090	2.1	20.6
6	Web search thesis	Yes	15 510	287.2	126	411	114	0.7	2.1	0	62	25	0.2	0.5
7	Search API thesis	—	45 903	850.1	620	941	2582	5.6	47.8	0	655	955	2.1	18.0
8	Search API thesis	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
9	Messerschmitt TG500 for sale	—	39 000	722.2	532	967	1097	2.8	20.3	0	419	344	0.9	6.5
10	Messerschmitt TG500 for sale	Yes	13 421	248.5	150	291	24	0.2	0.4	0	14	0	0.0	0.0
11	winds of winter	—	52 565	973.4	883	1000	2668	5.1	49.4	0	915	810	1.5	15.3
12	winds of winter	Yes	42 035	778.4	596	905	1717	4.1	31.8	0	481	635	1.5	12.0
13	promise of spring	—	52 698	975.9	863	1000	2685	5.1	49.7	0	931	906	1.7	17.1
14	promise of spring	Yes	44 616	826.2	455	912	2492	5.6	46.1	0	777	931	2.1	17.6
15	terry pratchett	—	45 437	841.4	388	995	2401	5.3	44.5	0	717	777	1.7	14.7
16	terry pratchett	Yes	41 235	763.6	586	963	1752	4.2	32.4	0	511	660	1.6	12.5
17	liverpool leads efl	—	51 214	948.4	866	999	3614	7.1	66.9	0	850	1160	2.3	21.9
18	liverpool leads efl	Yes	33 504	620.4	114	901	277	0.8	5.1	0	60	78	0.2	1.5
19	hillary clinton e-mail fbi	—	50 729	939.4	0	1000	2576	5.1	47.7	0	903	824	1.6	15.5
20	hillary clinton e-mail fbi	Yes	27 647	512.0	431	620	68	0.2	1.3	0	26	28	0.1	0.5
21	macbook pro 2016 touch bar problems	—	40 435	748.8	0	900	3099	7.7	57.4	0	617	1290	3.2	24.3
22	macbook pro 2016 touch bar problems	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
23	apple stock price	—	39 449	730.5	393	983	1736	4.4	32.1	0	586	524	1.3	9.9
24	apple stock price	Yes	36 366	673.4	264	1000	1098	3.0	20.3	0	194	482	1.3	9.1
25	samsung note 8 release date	—	47 934	887.7	0	995	2893	6.0	53.6	0	759	973	2.0	18.4
26	samsung note 8 release date	Yes	42 507	787.2	235	946	181	0.4	3.4	0	29	62	0.1	1.2
27	google self driving car	—	51 224	948.6	851	997	2540	5.0	47.0	0	882	752	1.5	14.2
28	google self driving car	Yes	40 914	757.7	614	906	1669	4.1	30.9	0	580	609	1.5	11.5
29	mobile application health sensor data	—	53 117	933.6	878	1000	2894	5.4	53.6	0	910	846	1.6	16.0
30	mobile application health sensor data	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
31	mobile phone body area network	—	52 870	979.1	882	1000	2206	4.2	40.9	0	876	510	1.0	9.6
32	mobile phone body area network	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
33	mobile phone sensor research health	—	52 996	981.4	825	1000	3019	5.7	55.9	0	932	823	1.6	15.5
34	mobile phone sensor research health	Yes	0	0.0	0	0	0	0.0	0.0	0	0	0	0.0	0.0
35	forest fairytales	—	52 958	980.7	877	1000	2542	4.8	47.1	0	930	832	1.6	15.7
36	forest fairytales	Yes	28 986	536.8	0	934	704	2.4	13.0	0	247	248	0.9	4.7
37	tudor politics	—	52 314	968.8	815	1000	2894	5.5	53.6	0	905	1044	2.0	19.7
38	tudor politics	Yes	39 589	733.1	526	966	1278	3.2	23.7	0	535	338	0.9	6.4
39	jazz poetry	—	51 367	951.2	0	1000	2640	5.1	48.9	0	940	802	1.6	15.1
40	jazz poetry	Yes	42 083	779.3	574	951	2102	5.0	38.9	0	663	791	1.9	14.9
	Sum		1 444 330				65 475					21 779		

Table A.3: Totals and averages for query 1 - 40, for the 54 day run period from day 18.