

# K3100-04

## (1) Evaluation Result without Considering Links

Summary Statistics	
Run ID	K3100-04
Subtask	II-A1
Relevant Level	Level1 (H or A)
Query Method	automatic
Topic Part	DISC only
Contents/Links Used	contents&links
Number of Topics	45
Total number of documents over all topics	
Retrieved:	29434
Relevant:	521
Rel-ret:	280

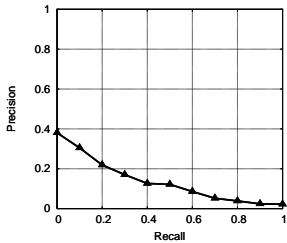
Summary Statistics	
Run ID	K3100-04
Subtask	II-A1
Relevant Level	Level2 (H, A or B)
Query Method	automatic
Topic Part	DISC only
Contents/Links Used	contents&links
Number of Topics	46
Total number of documents over all topics	
Retrieved:	30434
Relevant:	812
Rel-ret:	505

Recall Level Precision Averages	
Recall	Precision
0.00	0.3803
0.10	0.3045
0.20	0.2189
0.30	0.1701
0.40	0.1260
0.50	0.1215
0.60	0.0856
0.70	0.0517
0.80	0.0383
0.90	0.0237
1.00	0.0220
Average precision over all relevant documents	
non-interpolated	0.1267

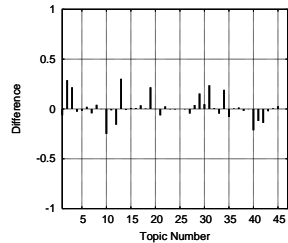
Document Level Averages	
	Precision
At 5 docs	0.1600
At 10 docs	0.1222
At 15 docs	0.1037
At 20 docs	0.0889
At 30 docs	0.0741
At 100 docs	0.0389
At 200 docs	0.0246
At 500 docs	0.0117
At 1000 docs	0.0062
R-Precision (precision after R docs retrieved (where R is the number of relevant documnets))	
Exact	0.1203

Recall Level Precision Averages	
Recall	Precision
0.00	0.4991
0.10	0.3816
0.20	0.2887
0.30	0.1970
0.40	0.1549
0.50	0.1358
0.60	0.1096
0.70	0.0623
0.80	0.0505
0.90	0.0352
1.00	0.0240
Average precision over all relevant documents	
non-interpolated	0.1563

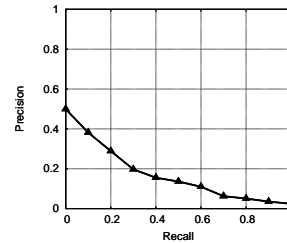
Document Level Averages	
	Precision
At 5 docs	0.2261
At 10 docs	0.1739
At 15 docs	0.1493
At 20 docs	0.1337
At 30 docs	0.1152
At 100 docs	0.0637
At 200 docs	0.0416
At 500 docs	0.0205
At 1000 docs	0.0110
R-Precision (precision after R docs retrieved (where R is the number of relevant documnets))	
Exact	0.1650



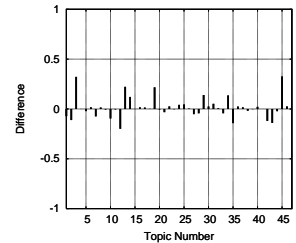
Recall-Precision Curve



Difference from Median in Average Precision per Topic (over 12 'automatic' and 'DISC only' runs)\*

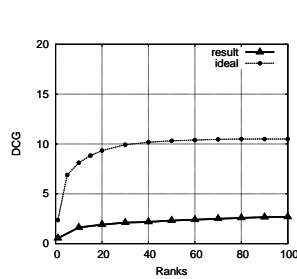


Recall-Precision Curve



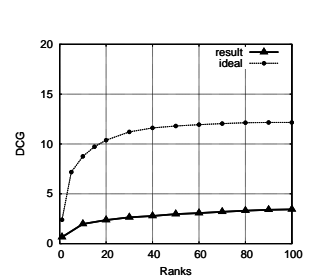
Difference from Median in Average Precision per Topic (over 12 'automatic' and 'DISC only' runs)\*\*

Discounted ( $\log_2$ ) Cumulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 0$ )	
	DCG
At 1 doc	0.5435
At 10 docs	1.6137
At 20 docs	1.9059
At 30 docs	2.1072
At 40 docs	2.1746
At 50 docs	2.3163
At 60 docs	2.4036
At 70 docs	2.4939
At 80 docs	2.5778
At 90 docs	2.6523
At 100 docs	2.6787
At 200 docs	2.9729
At 500 docs	3.2105
At 1000 docs	3.2887



Discounted Cumulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 0$ )

Discounted ( $\log_2$ ) Cumulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 1$ )	
	DCG
At 1 doc	0.6522
At 10 docs	1.9706
At 20 docs	2.3621
At 30 docs	2.6379
At 40 docs	2.7734
At 50 docs	2.9623
At 60 docs	3.0646
At 70 docs	3.1908
At 80 docs	3.3060
At 90 docs	3.3942
At 100 docs	3.4338
At 200 docs	3.8637
At 500 docs	4.2208
At 1000 docs	4.3383



Discounted Cumulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 1$ )

\* Topic Numbers: ( 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45 )  
 Topic IDs: ( 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 52, 53, 58, 59, 60, 61, 62, 63 )

\*\* Topic Numbers: ( 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 )  
 Topic IDs: ( 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47, 48, 49, 52, 53, 57, 58, 59, 60, 61, 62, 63 )

## (2) Evaluation Result with Considering Links

Summary Statistics	
Run ID	K3100-04
Subtask	II-A1
Relevant Level	Level1 (H or A)
Query Method	automatic
Topic Part	DISC only
Contents/Links Used	contents&links
Number of Topics	46
Total number of documents over all topics	
Retrieved:	30434
Relevant:	966
Rel-ret:	474

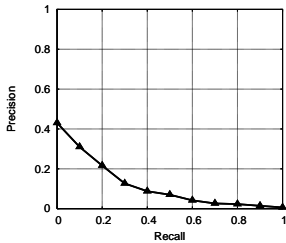
Summary Statistics	
Run ID	K3100-04
Subtask	II-A1
Relevant Level	Level2 (H, A or B)
Query Method	automatic
Topic Part	DISC only
Contents/Links Used	contents&links
Number of Topics	47
Total number of documents over all topics	
Retrieved:	31434
Relevant:	1381
Rel-ret:	748

Recall Level Precision Averages	
Recall	Precision
0.00	0.4304
0.10	0.3099
0.20	0.2162
0.30	0.1267
0.40	0.0876
0.50	0.0700
0.60	0.0418
0.70	0.0263
0.80	0.0230
0.90	0.0144
1.00	0.0062
Average precision over all relevant documents	
non-interpolated	0.1014

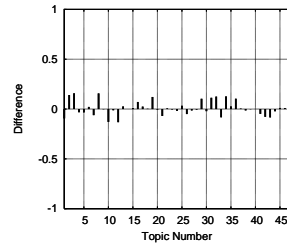
Document Level Averages	
	Precision
At 5 docs	0.2130
At 10 docs	0.1652
At 15 docs	0.1406
At 20 docs	0.1207
At 30 docs	0.1022
At 100 docs	0.0578
At 200 docs	0.0386
At 500 docs	0.0192
At 1000 docs	0.0103
R-Precision (precision after R docs retrieved (where R is the number of relevant documnets))	
Exact	0.1243

Recall Level Precision Averages	
Recall	Precision
0.00	0.5142
0.10	0.3309
0.20	0.2261
0.30	0.1510
0.40	0.1087
0.50	0.0755
0.60	0.0585
0.70	0.0379
0.80	0.0309
0.90	0.0166
1.00	0.0054
Average precision over all relevant documents	
non-interpolated	0.1192

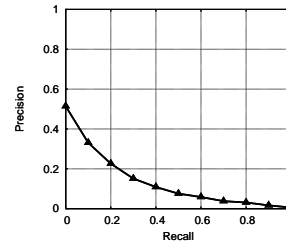
Document Level Averages	
	Precision
At 5 docs	0.2766
At 10 docs	0.2191
At 15 docs	0.1872
At 20 docs	0.1660
At 30 docs	0.1426
At 100 docs	0.0868
At 200 docs	0.0587
At 500 docs	0.0297
At 1000 docs	0.0159
R-Precision (precision after R docs retrieved (where R is the number of relevant documnets))	
Exact	0.1463



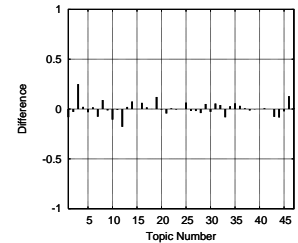
Recall-Precision Curve



Difference from Median in Average Precision per Topic (over 12 'automatic' and 'DISC only' runs)\*

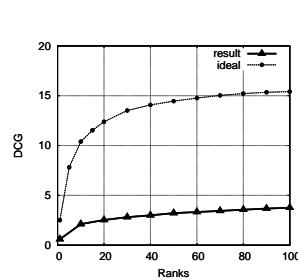


Recall-Precision Curve



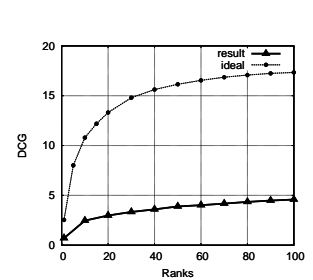
Difference from Median in Average Precision per Topic (over 12 'automatic' and 'DISC only' runs)\*\*

Discounted ( $\log_2$ ) Cummulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 1$ )	
	DCG
At 1 doc	0.5957
At 10 docs	2.1115
At 20 docs	2.5243
At 30 docs	2.8130
At 40 docs	2.9944
At 50 docs	3.2144
At 60 docs	3.3256
At 70 docs	3.4387
At 80 docs	3.5619
At 90 docs	3.6746
At 100 docs	3.7586
At 200 docs	4.3276
At 500 docs	4.8099
At 1000 docs	4.9564



Discounted Cummulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 0$ )

Discounted ( $\log_2$ ) Cummulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 1$ )	
	DCG
At 1 doc	0.7021
At 10 docs	2.4641
At 20 docs	2.9745
At 30 docs	3.3317
At 40 docs	3.5918
At 50 docs	3.8848
At 60 docs	4.0143
At 70 docs	4.1731
At 80 docs	4.3408
At 90 docs	4.4768
At 100 docs	4.5802
At 200 docs	5.3140
At 500 docs	5.9434
At 1000 docs	6.1327



Discounted Cummulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 1$ )

\* Topic Numbers: ( 1, 2, 3, 4, 5, 6, 7, 8, 9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46 )

Topic IDs: ( 8,10,11,12,13,14,15,16,17,18,19,20,22,23,24,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,46,47,48,49,52,53,56,58,59,60,61,62,63 )

\*\* Topic Numbers: ( 1, 2, 3, 4, 5, 6, 7, 8, 9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47 )

Topic IDs: ( 8,10,11,12,13,14,15,16,17,18,19,20,22,23,24,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,46,47,48,49,52,53,56,57,58,59,60,61,62,63 )