Terabyte Track results — Tsinghua University

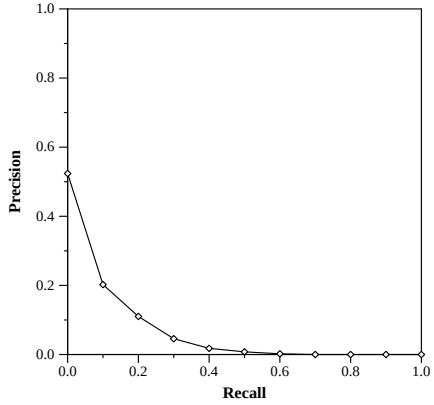
Summary Statistics	
Run ID	THUIRtb2
Run Type	automatic
Fields Used	title, description, narrative
Run Description	NoLinksUsed-AnchorTextUsed-DocStructUsed
Time to index	170 min.
Time to retrieve top 20 per query	18 sec.
System RAM	0.75 GB
Size of on-disk structures	16 GB
Approx system cost	$800
Number of Topics:	49
	Total number of documents over all topics
Retrieved:	49000
Relevant:	10617
Rel-ret:	1764

Recall Level Averages	
Recall	Precision
0.00	0.5237
0.10	0.2023
0.20	0.1102
0.30	0.0457
0.40	0.0179
0.50	0.0077
0.60	0.0021
0.70	0.0000
0.80	0.0000
0.90	0.0000
1.00	0.0000

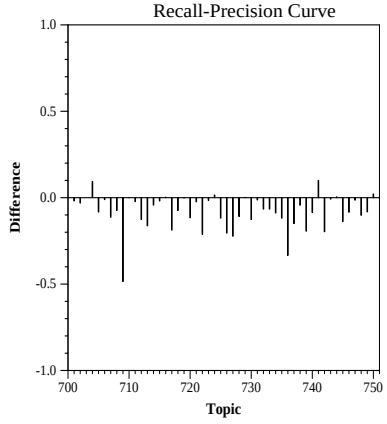
Mean average precision	
non-interpolated	0.0557

Document Level Averages	
	Precision
At 5 docs	0.3469
At 10 docs	0.3061
At 15 docs	0.2803
At 20 docs	0.2663
At 30 docs	0.2340
At 100 docs	0.1373
At 200 docs	0.0976
At 500 docs	0.0554
At 1000 docs	0.0360

R-Precision: precision after R (number relevant) documents retrieved	
Exact	0.0983



Recall-Precision Curve



Difference from Median Average Precision per Topic

Terabyte Track results — Tsinghua University

Summary Statistics	
Run ID	THUIRtb3
Run Type	automatic
Fields Used	title
Run Description	NoLinksUsed-AnchorTextUsed-NoDocStructUsed
Time to index	1024 min.
Time to retrieve top 20 per query	9 sec.
System RAM	1.5 GB
Size of on-disk structures	121 GB
Approx system cost	$1000

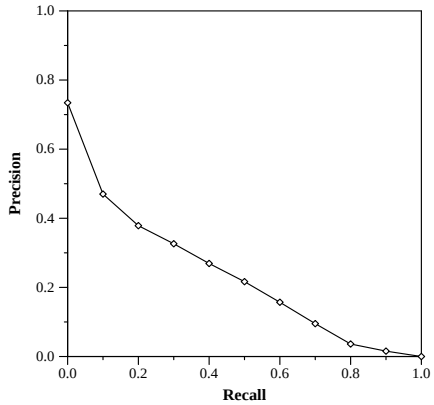
Number of Topics:	49
	Total number of documents over all topics
Retrieved:	49000
Relevant:	10617
Rel-ret:	5843

Recall Level Averages	
Recall	Precision
0.00	0.7342
0.10	0.4700
0.20	0.3786
0.30	0.3263
0.40	0.2692
0.50	0.2166
0.60	0.1570
0.70	0.0951
0.80	0.0362
0.90	0.0156
1.00	0.0000

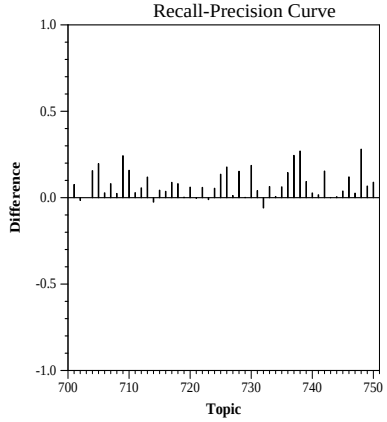
Mean average precision	
non-interpolated	0.2198

Document Level Averages	
	Precision
At 5 docs	0.4857
At 10 docs	0.4714
At 15 docs	0.4707
At 20 docs	0.4663
At 30 docs	0.4503
At 100 docs	0.3594
At 200 docs	0.2917
At 500 docs	0.1841
At 1000 docs	0.1192

R-Precision: precision after R (number relevant) documents retrieved	
Exact	0.2968



Recall-Precision Curve



Difference from Median Average Precision per Topic

Terabyte Track results — Tsinghua University

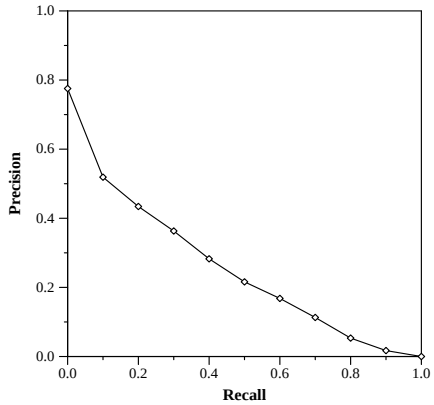
Summary Statistics	
Run ID	THUIRtb4
Run Type	automatic
Fields Used	title, description, narrative
Run Description	NoLinksUsed-AnchorTextUsed-NoDocStructUsed
Time to index	1024 min.
Time to retrieve top 20 per query	55 sec.
System RAM	1.5 GB
Size of on-disk structures	121 GB
Approx system cost	$1000
Number of Topics:	49
	Total number of documents over all topics
Retrieved:	49000
Relevant:	10617
Rel-ret:	5690

Recall Level Averages	
Recall	Precision
0.00	0.7754
0.10	0.5189
0.20	0.4341
0.30	0.3632
0.40	0.2827
0.50	0.2160
0.60	0.1681
0.70	0.1130
0.80	0.0533
0.90	0.0172
1.00	0.0000

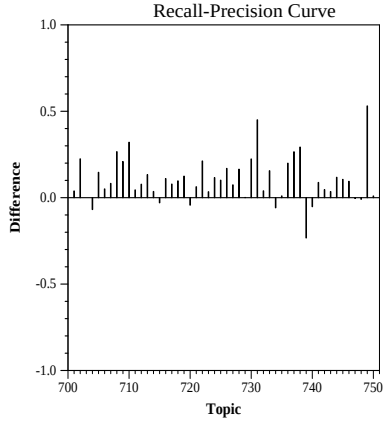
Mean average precision	
non-interpolated	0.2453

Document Level Averages	
	Precision
At 5 docs	0.5796
At 10 docs	0.5327
At 15 docs	0.5279
At 20 docs	0.5051
At 30 docs	0.4803
At 100 docs	0.3741
At 200 docs	0.2882
At 500 docs	0.1825
At 1000 docs	0.1161

R-Precision: precision after R (number relevant) documents retrieved	
Exact	0.3049



Recall-Precision Curve



Difference from Median Average Precision per Topic

Terabyte Track results — Tsinghua University

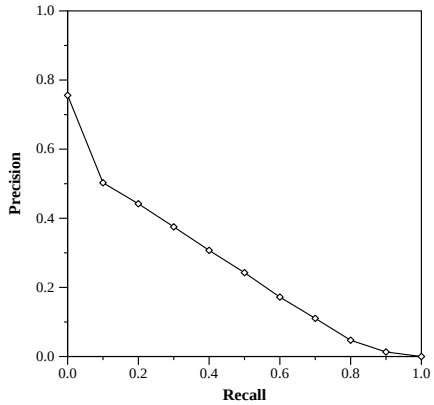
Summary Statistics	
Run ID	THUIRtb5
Run Type	automatic
Fields Used	title
Run Description	NoLinksUsed-NoAnchorTextUsed-NoDocStructUsed
Time to index	1920 min.
Time to retrieve top 20 per query	15 sec.
System RAM	0.5 GB
Size of on-disk structures	41.5 GB
Approx system cost	$750
Number of Topics:	49
	Total number of documents over all topics
Retrieved:	49000
Relevant:	10617
Rel-ret:	6134

Recall Level Averages	
Recall	Precision
0.00	0.7557
0.10	0.5027
0.20	0.4420
0.30	0.3752
0.40	0.3072
0.50	0.2426
0.60	0.1720
0.70	0.1104
0.80	0.0471
0.90	0.0132
1.00	0.0000

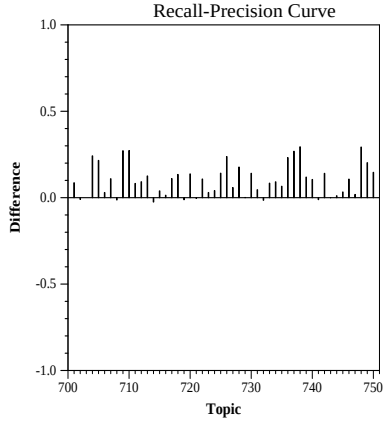
Mean average precision	
non-interpolated	0.2437

Document Level Averages	
	Precision
At 5 docs	0.4816
At 10 docs	0.4755
At 15 docs	0.4680
At 20 docs	0.4653
At 30 docs	0.4571
At 100 docs	0.3924
At 200 docs	0.3166
At 500 docs	0.1966
At 1000 docs	0.1252

R-Precision: precision after R (number relevant) documents retrieved	
Exact	0.3181



Recall-Precision Curve



Difference from Median Average Precision per Topic

Terabyte Track results — Tsinghua University

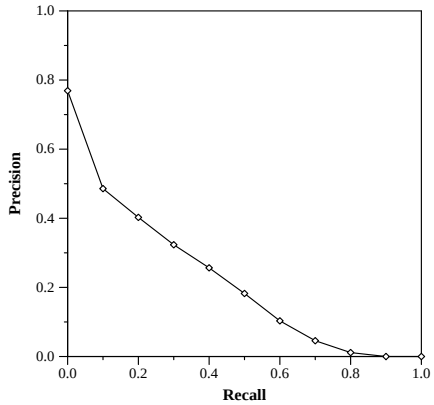
Summary Statistics	
Run ID	THUIRtb6
Run Type	automatic
Fields Used	title
Run Description	NoLinksUsed-NoAnchorTextUsed-NoDocStructUsed
Time to index	1920 min.
Time to retrieve top 20 per query	16 sec.
System RAM	0.5 GB
Size of on-disk structures	41.5 GB
Approx system cost	$750
Number of Topics:	49
	Total number of documents over all topics
Retrieved:	49000
Relevant:	10617
Rel-ret:	5092

Recall Level Averages	
Recall	Precision
0.00	0.7691
0.10	0.4858
0.20	0.4026
0.30	0.3234
0.40	0.2566
0.50	0.1824
0.60	0.1032
0.70	0.0457
0.80	0.0114
0.90	0.0000
1.00	0.0000

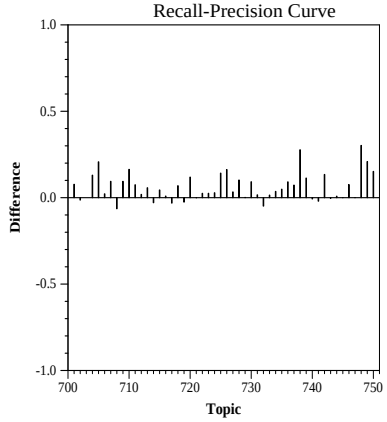
Mean average precision	
non-interpolated	0.2036

Document Level Averages	
	Precision
At 5 docs	0.5020
At 10 docs	0.4735
At 15 docs	0.4653
At 20 docs	0.4561
At 30 docs	0.4490
At 100 docs	0.3714
At 200 docs	0.2861
At 500 docs	0.1704
At 1000 docs	0.1039

R-Precision: precision after R (number relevant) documents retrieved	
Exact	0.2864



Recall-Precision Curve



Difference from Median Average Precision per Topic