

Terabyte Track results — Microsoft Research Asia

Summary Statistics	
Run ID	MSRAt4
Run Type	automatic
Fields Used	title
Run Description	NoLinksUsed-AnchorTextUsed- DocStructUsed
Time to index	700 min.
Time to retrieve top 20 per query	1 sec.
System RAM	4 GB
Size of on-disk structures	54 GB
Approx system cost	\$28200
Number of Topics:	49
Total number of documents over all topics	
Retrieved:	49000
Relevant:	10617
Rel-ret:	4932

Recall Level Averages	
Recall	Precision
0.00	0.7455
0.10	0.4658
0.20	0.3856
0.30	0.3004
0.40	0.2321
0.50	0.1633
0.60	0.0617
0.70	0.0158
0.80	0.0019
0.90	0.0000
1.00	0.0000
Mean average precision	
non-interpolated	0.1877

Document Level Averages	
	Precision
At 5 docs	0.4898
At 10 docs	0.4959
At 15 docs	0.4503
At 20 docs	0.4337
At 30 docs	0.4122
At 100 docs	0.3469
At 200 docs	0.2741
At 500 docs	0.1661
At 1000 docs	0.1007
R-Precision: precision after R (number relevant) docu- ments retrieved	
Exact	0.2741

