

Round 1 results — Run RMITBM1 submitted from RMITB

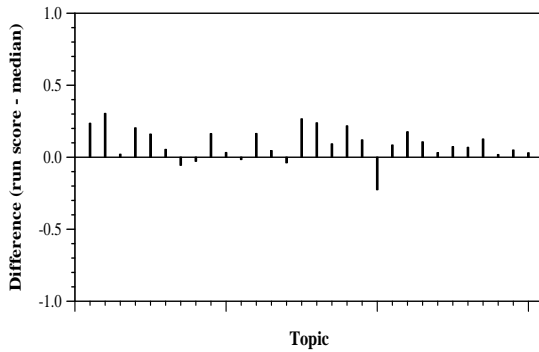
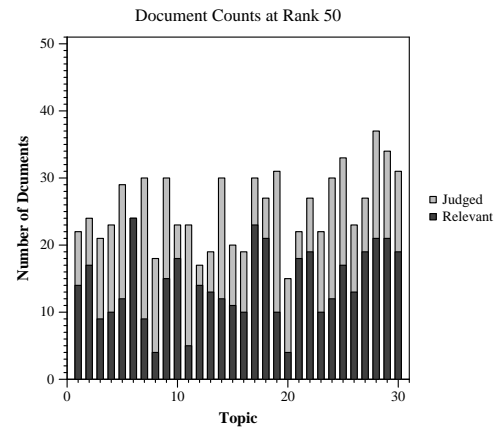
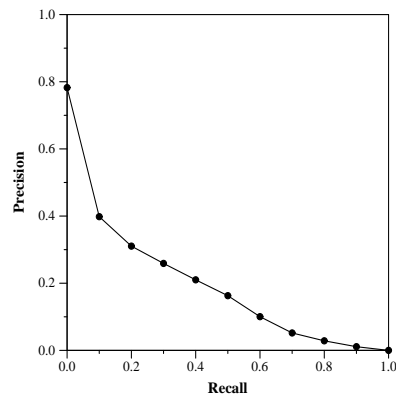
Run Description

10 human query variations for each topic narrative were produced. Each of these queries were fed into 16 Terrier retrieval models with no query expansion applied, fused with CombSUM. These relevance scores were further minmaxed between 0 and 1, to be linearly combined with an exponential decay function for recency of publication based on the cord metadata, where docs older than 4 months are not given a buff in score.

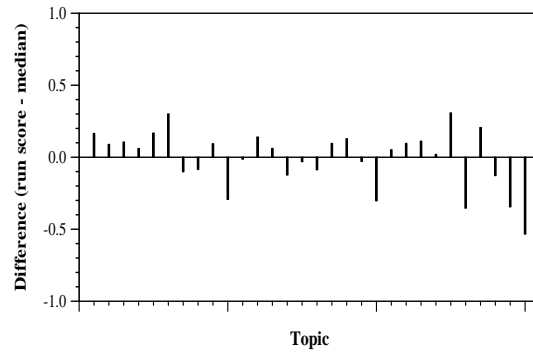
Summary Statistics	
Run ID	RMITBM1
Topic type	manual
Contributed to judgment sets?	yes

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	1380
MAP	0.1780
Mean Bpref	0.4162
Mean NDCG@10	0.3511

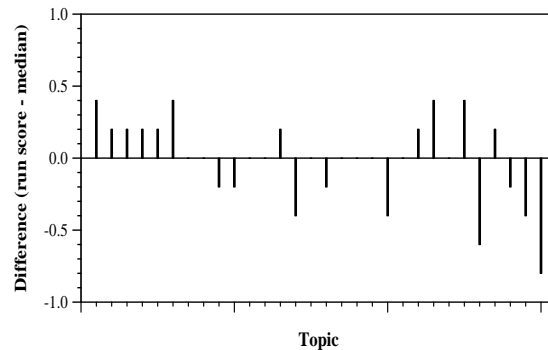
Document Level Averages	
	Precision
At 5 docs	0.4733
At 10 docs	0.3867
At 15 docs	0.3444
At 20 docs	0.3200
At 30 docs	0.2989
R-Precision	
Exact	0.2509



Per-topic difference from median bpref for all Round 1 runs



Per-topic difference from median NDCG@10 for all Round 1 runs



Per-topic difference from median P@5 for all Round 1 runs