

Round 1 results — Run RMITBFuseM2 submitted from RMITB

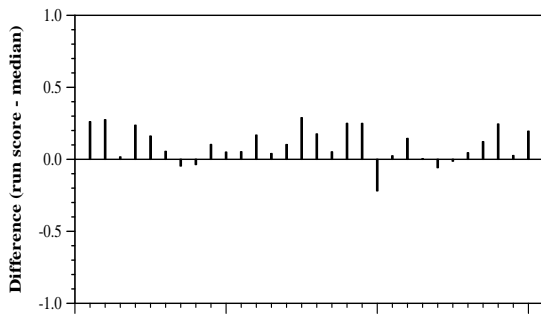
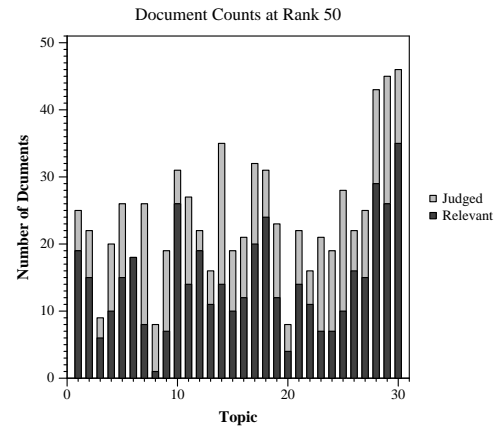
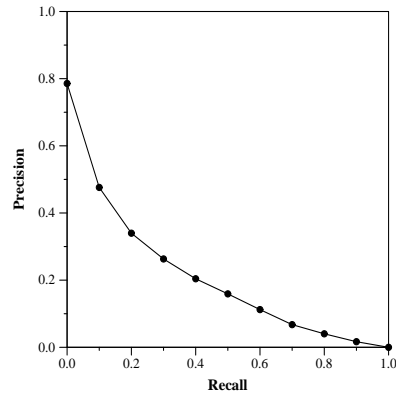
Run Description

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. No buffs were applied to preference recently published articles.

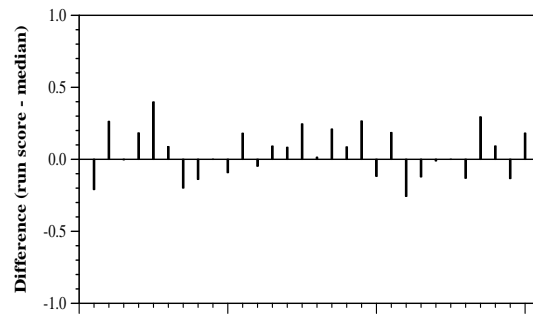
Summary Statistics	
Run ID	RMITBFuseM2
Topic type	manual
Contributed to judgment sets?	no

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	1380
MAP	0.1968
Mean Bpref	0.4247
Mean NDCG@10	0.4053

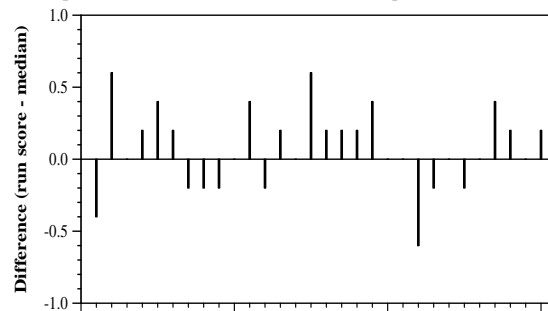
Document Level Averages	
	Precision
At 5 docs	0.5600
At 10 docs	0.4767
At 15 docs	0.4267
At 20 docs	0.4067
At 30 docs	0.3622
R-Precision	
Exact	0.2600



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