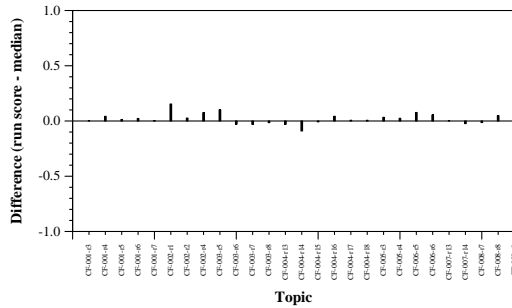


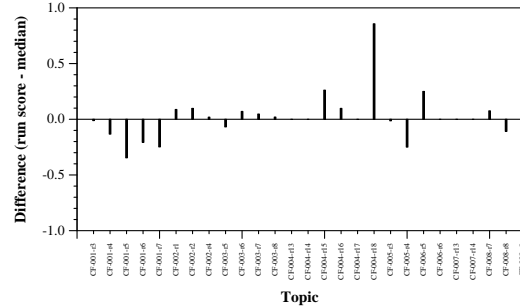
Summary Statistics	
Run ID	ColBERT_ILP
Run type	Automatic
TREC-IS labels?	no
Streams	Twitter; Reddit; Web news
Extractive or abstractive	extractive

Manual measure means	isolated	accumulative
Comprehensiveness	0.1840	0.1332
Redundancy Ratio	0.1945	0.4049
Assessed %	0.5693	0.5693
Matched %	0.3012	0.3012
Skipped %	0.0000	0.0000
Irrelevant %	0.2633	0.2633
Unmatched %	0.4355	0.4355

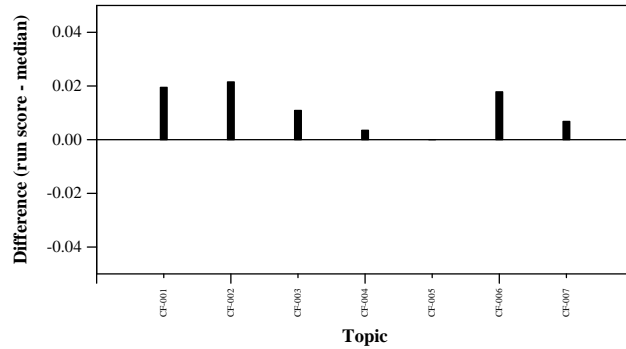
Automatic measure means	ICS	Wikipedia	NIST
ROUGE-1 F1	0.0051	0.0022	0.0098
ROUGE-2 F1	0.0011	0.0007	0.0031
ROUGE-L F1	0.0017	0.0011	0.0026
ROUGE-Lsum F1	0.0050	0.0021	0.0092
BERTscore F1	0.0100	0.0121	0.0121



Per-topic difference from median accumulative comprehensiveness



Per-topic difference from median accumulative redundancy ratio



Per-topic difference from median ROUGE-2 to ICS