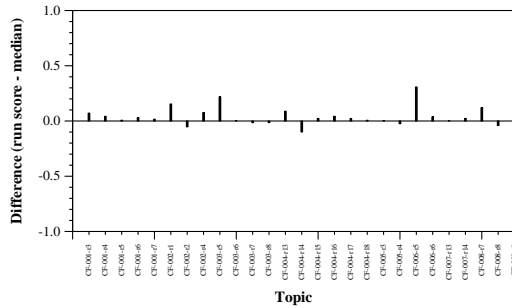


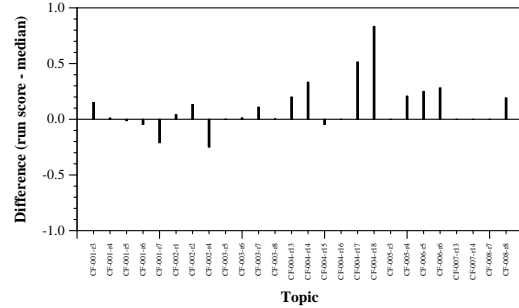
Summary Statistics	
Run ID	BM25-QAasnq_ILP
Run type	Automatic
TREC-IS labels?	no
Streams	Twitter; Reddit; Web news
Extractive or abstractive	extractive

Manual measure means	isolated	accumulative
Comprehensiveness	0.2065	0.1535
Redundancy Ratio	0.2300	0.4868
Assessed %	0.5710	0.5710
Matched %	0.3816	0.3816
Skipped %	0.0044	0.0044
Irrelevant %	0.1975	0.1975
Unmatched %	0.4164	0.4164

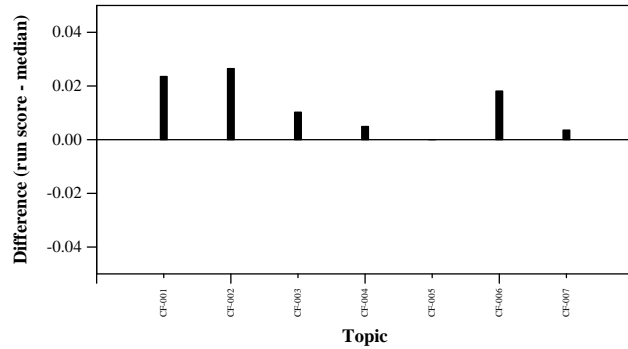
Automatic measure means	ICS	Wikipedia	NIST
ROUGE-1 F1	0.0051	0.0024	0.0099
ROUGE-2 F1	0.0011	0.0008	0.0033
ROUGE-L F1	0.0018	0.0012	0.0027
ROUGE-Lsum F1	0.0050	0.0022	0.0094
BERTscore F1	0.0099	0.0125	0.0125



Per-topic difference from median accumulative comprehensiveness



Per-topic difference from median accumulative redundancy ratio



Per-topic difference from median ROUGE-2 to ICS