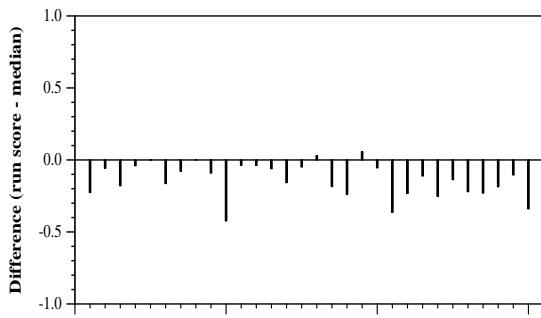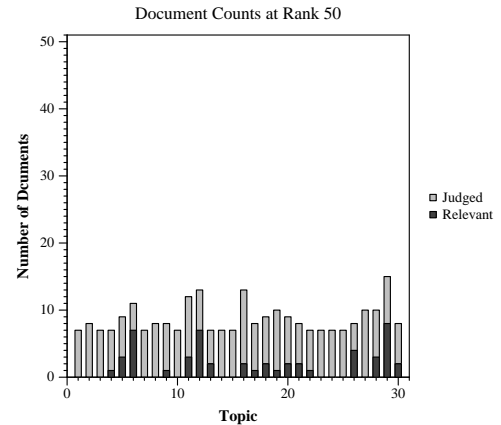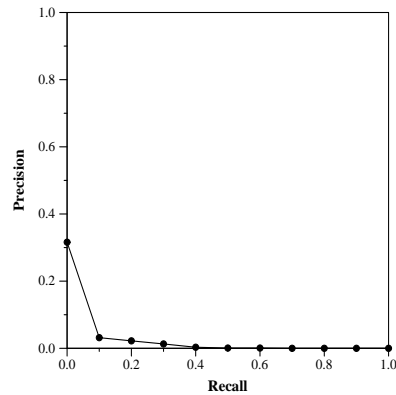## Run Description

The submission was fully automatic. Topic <query> and <question> elements were Stop Word and POS tag filtered (using nltk) to produce a keyword list, that was expanded by querying a Word2Vec model (gensim) of the CORD-19 dataset for 10 most related terms. Expanded keyword list was encoded as a vector in a doc2vec model. This model was trained using gensim lib after pre-processing for phrase extraction with spacy, textacy and Abner NER for multi-word named entity. The 1000 most related doc were returned.
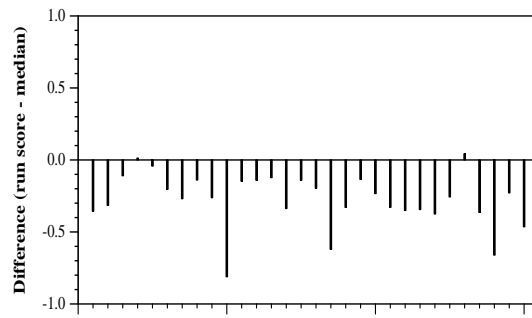
Summary Statistics	
Run ID	TMACC_SeTA_baseline
Topic type	automatic
Contributed to judgment sets?	yes

Overall measures	
Number of topics	30
Total number retrieved	30000
Total relevant	2352
Total relevant retrieved	602
MAP	0.0132
Mean Bpref	0.1870
Mean NDCG@10	0.0847

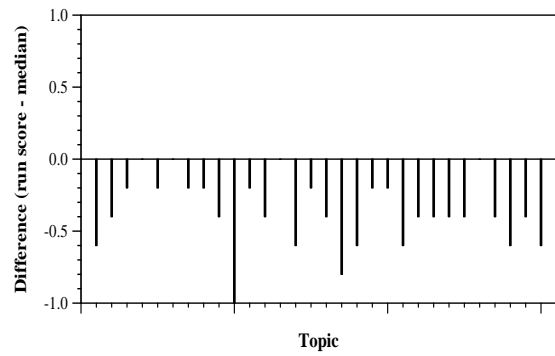| Document Level Averages	
	Precision
At 5 docs	0.1200
At 10 docs	0.0833
At 15 docs	0.0644
At 20 docs	0.0550
At 30 docs	0.0444
R-Precision	
Exact	0.0302

Document Counts at Rank 50

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