Podcasts Track results — (udel_wang_zheng) University of Delaware
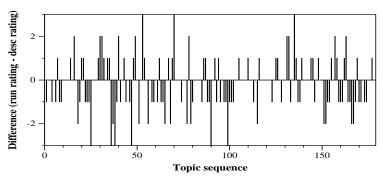
Summary Statistics	
Run ID	udel_wang_zheng1
Task	Summarization
Run type	automatic
Summarization model	Linguistic model; Supervised neural or deep learning model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	13 (7.3%)
Summaries rated "good"	39 (21.8%)
Summaries rated "fair"	83 (46.4%)
Summaries rated "bad"	44 (24.6%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2387 (95%-conf.int. 0.22744 - 0.24958)
ROUGE_L R score	0.1607 (95%-conf.int. 0.15292 - 0.16882)
ROUGE_L F score	0.1630 (95%-conf.int. 0.15603 - 0.17008)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	59.8%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	34.1%
Q3: Does the summary include the main topic(s) of the podcast?	69.8%
Q4: Does the summary tell you anything about the format of the podcast; e.g. whether it's an interview, whether it's a chat between friends, a monologue, etc?	50.8%
Q5: Does the summary give you mode context on the title of the podcast?	54.2%
Q6: Does the summary contain redundant information?	5.0%
Q7: Is the summary written in good English?	73.7%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	50.3%

Per-topic difference in rating from episode description