

Podcasts Track results — (TREMA-UNH) University of New Hampshire

Summary Statistics	
Run ID	unhtrema4
Task	Summarization
Run type	automatic
Summarization model	Supervised neural or deep learning model
Summarization method	Other
Number of summaries	179

Overall measures	
Summaries rated "excellent"	1 (0.6%)
Summaries rated "good"	0 (0.0%)
Summaries rated "fair"	4 (2.2%)
Summaries rated "bad"	174 (97.2%)

ROUGE score compared to episode description	
ROUGE_L P score	0.0689 (95%-conf.int. 0.06554 - 0.07243)
ROUGE_L R score	0.1795 (95%-conf.int. 0.17451 - 0.18514)
ROUGE_L F score	0.0853 (95%-conf.int. 0.08306 - 0.08765)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	1.1%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	0.6%
Q3: Does the summary include the main topic(s) of the podcast?	9.5%
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?	2.2%
Q5: Does the summary give you more context on the title of the podcast?	5.6%
Q6: Does the summary contain redundant information?	9.5%
Q7: Is the summary written in good English?	1.1%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	0.6%

