

Podcasts Track results — (podcast_baselines) Baseline runs for Podcast track

Summary Statistics	
Run ID	onemin
Task	Summarization
Run type	automatic
Summarization model	Other
Summarization method	Extractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	5 (2.8%)
Summaries rated "good"	25 (14.0%)
Summaries rated "fair"	96 (53.6%)
Summaries rated "bad"	53 (29.6%)

ROUGE score compared to episode description	
ROUGE_L P score	0.0871 (95%-conf.int. 0.08265 - 0.09204)
ROUGE_L R score	0.2820 (95%-conf.int. 0.27209 - 0.29250)
ROUGE_L F score	0.1157 (95%-conf.int. 0.11122 - 0.12072)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	69.8%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	44.1%
Q3: Does the summary include the main topic(s) of the podcast?	65.9%
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?	52.5%
Q5: Does the summary give you more context on the title of the podcast?	62.6%
Q6: Does the summary contain redundant information?	24.0%
Q7: Is the summary written in good English?	45.3%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	9.5%

