

## Podcasts Track results — (spotify) Spotify

Summary Statistics	
Run ID	categoryaware2
Task	Summarization
Run type	automatic
Summarization 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"	32 (17.9%)
Summaries rated "good"	39 (21.8%)
Summaries rated "fair"	77 (43.0%)
Summaries rated "bad"	31 (17.3%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2575 (95%-conf.int. 0.24471 - 0.26970)
ROUGE_L R score	0.1986 (95%-conf.int. 0.18883 - 0.20844)
ROUGE_L F score	0.1842 (95%-conf.int. 0.17559 - 0.19256)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	55.3%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	40.8%
Q3: Does the summary include the main topic(s) of the podcast?	77.7%
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?	57.0%
Q5: Does the summary give you more context on the title of the podcast?	65.9%
Q6: Does the summary contain redundant information?	6.7%
Q7: Is the summary written in good English?	88.8%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	70.4%

