

## EXHIBIT A

### EXHIBIT A: DESCRIPTION OF THE DATASET

AP16-OL7 Multilingual Database			Training/Development Set			Testing Set		
Code	Language	Channel	No. of Speaker	Utterance/Speaker	Total Utterances	Nos. of Speaker	Utterance/Speaker	Total Utterances
King-ASR-L-058	Cantonese in China Mainland and Hongkong	Mobile	18	320	5760	6	320	1920
King-ASR-L-216	Mandarin in China	Mobile	18	300	5400	6	300	1800
King-ASR-L-224	Indonesian in Indonesia	Mobile	18	320	5760	6	320	1920
King-ASR-L-269	Japanese in Japan	Mobile	18	320	5760	6	320	1920
King-ASR-L-124	Russian in Russia	Mobile	18	300	5400	6	300	1800
King-ASR-L-137	Korean in Korea	Mobile	18	300	5400	6	300	1800
King-ASR-L-225	Vietnamese in Vietnam	Mobile	18	365	6570	6	365	2190

Note

- \* Male and Female are balanced;
- \* Original scripts are available;
- \* Lexicon are available;
- \* There might be a minor discrepancy with the numbers of total utterances;