

What Is Claimed Is:

1. A speech processing system including one or more processors and one or more memories configured to perform operations comprising:
 - capturing one or more speech signals, wherein each of the one or more speech signals comprises at least one dialogue uttered by a user;
 - extracting dialogues from the one or more speech signals;
 - identifying frequently uttered dialogues over a period of time, wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times during the period of time more than other dialogues uttered by the user during the period of time;
 - generating a local language model and a local acoustic model based on, at least in part, the frequently uttered dialogues; and
 - processing the one or more speech signals based on, at least in part, the local language model and the local acoustic model.
2. The speech processing system according to claim 1, wherein the extracting dialogues from the one or more speech signals further comprises implementing a remote language model and a remote acoustic model for extracting the dialogues from the one or more speech signals, and wherein the local language model and the local acoustic model are subsets of the remote language model and the remote acoustic model, respectively.
3. The speech processing system according to claim 2, wherein the local language model and the local acoustic model are executed on a user device and the remote language model and the remote acoustic model are executed remotely from the user device on at least one server.

4. The speech processing system according to claim 1, wherein the operations further comprising categorizing the dialogues into different domains.
5. The speech processing system according to claim 4, wherein the operations further comprising after categorizing the dialogues into different domains, ranking the categorized dialogues in each of the domains based on, at least in part, a frequency of utterance of each of the dialogues by the user over the period of time, such that the dialogues with a higher frequency of utterance are ranked higher as compared to the dialogues with a relatively lower frequency of utterance in the corresponding domain.
6. The speech processing system according to claim 5, wherein the identifying frequently uttered dialogues further comprises identifying a predefined number of dialogues with high rankings in each of the domains as the frequently uttered dialogues.
7. The speech processing system according to claim 6, wherein the predefined number of dialogues are determined based on, at least in part, one or more of a memory and a processor of a user device.
8. The speech processing system according to claim 5, wherein the domains include one or more of food, entertainment, sports, scheduling, sales inquiry, and automation commands.

9. A local speech processing system including one or more processors and one or more memories configured to perform operations comprising:
 - capturing one or more speech signals, wherein each of the one or more speech signals includes at least one dialogue uttered by a user;
 - receiving a local language model and a local acoustic model from a remote automatic speech recognition (ASR) engine;
 - storing the local language model and the local acoustic model, wherein the local language model and the local acoustic model is based on, at least in part, frequently uttered dialogues over a period of time, wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times during the period of time more than other dialogues uttered by the user during the period of time; and
 - processing the one or more speech signals based on, at least in part, the language model and the acoustic model.
10. The local speech processing system according to claim 9, wherein the operations further comprising associating the local language model and the local acoustic model with a user profile of the user.
11. The local speech processing system according to claim 10, wherein the storing the local language model and the local acoustic further comprises storing the local language model and the local acoustic model in a memory of a user device linked with the user profile of the user.
12. The local speech processing system according to claim 9, wherein the processing the one or more speech signals further comprises executing a local

ASR engine on a user device to process the one or more speech signals based on, at least in part, the local language model and the local acoustic model.

13. A remote speech processing system including one or more processors and one or more memories configured to perform operations comprising:

- receiving one or more speech signals from a user device, wherein each of the one or more speech signals includes at least one dialogue uttered by a user;

- extracting dialogues from the one or more speech signals;

- identifying frequently uttered dialogues over a period of time wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times during the period of time more than other dialogues uttered by the user during the period of time;

- generating a local language model and a local acoustic model for the user based on, at least in part, the frequently uttered dialogues by the user; and

- sending the local language model and the local acoustic model to the user device to be used in processing of the one or more speech signals.

14. The remote speech processing system according to claim 13, wherein the operations further comprising associating the local language model and the local acoustic model with a user profile of the user.

15. The remote speech processing system according to claim 13, wherein the extracting dialogues from the one or more speech signals further comprises implementing a remote ASR engine including a remote language model and a remote acoustic model for extracting the dialogues from the one or more speech signals, and wherein the local language model and the local acoustic model are

subsets of the remote language model and the remote acoustic model, respectively.

16. The remote speech processing system according to claim 15, wherein the remote ASR engine is executed remotely from the user device on at least one server.
17. The remote speech processing system according to claim 15, wherein at least a portion of the remote ASR engine is executed locally on the user device.
18. The remote speech processing system according to claim 13, wherein the local language model and the local acoustic model are used by a local ASR engine on the user device.
19. The remote speech processing system according to claim 13, wherein the operations further comprising categorizing the dialogues into different domains.
20. The remote speech processing system according to claim 19, wherein the operations further comprising after categorizing the dialogues into different domains, ranking the categorized dialogues in each of the domains based on, at least in part, a frequency of utterance of each of the dialogues by the user over the period of time, such that the dialogues with a higher frequency of utterance are ranked higher as compared to the dialogues with a relatively lower frequency of utterance in the corresponding domain.

21. The remote speech processing system according to claim 20, wherein the identifying frequently uttered dialogues further comprises identifying a predefined number of dialogues with high rankings in each of the domains as the frequently uttered dialogues.
22. The remote speech processing system according to claim 21, wherein the predefined number of dialogues are determined based on, at least in part, one or more of a memory and a processor of the user device.
23. The speech processing system according to claim 19, wherein the domains include one or more of food, entertainment, sports, scheduling, sales inquiry, and automation commands.
24. A computer implemented method for processing speech, comprising:
 - receiving one or more speech signals, wherein each of the one or more speech signals comprises at least one dialogue uttered by a user;
 - extracting dialogues from the one or more speech signals;
 - identifying frequently uttered dialogues over a period of time, wherein the frequently uttered dialogues are a set of dialogues that are uttered by the user a number of times during the period of time more than other dialogues uttered by the user during the period of time;
 - generating a local language model and a local acoustic model based on, at least in part, the identified frequently uttered dialogues; and
 - processing the one or more speech signals from the user based on, at least in part, the local language model and the local acoustic model.

25. The method according to claim 24 further comprising implementing a remote language model and a remote acoustic model for generating the local language model and the local acoustic model, wherein the local language model and the local acoustic model are subsets of the remote language model and the remote acoustic model, respectively.
26. The method according to claim 24 further comprising storing the local language model and the local acoustic model for later retrieval.
27. The method according to claim 24 further comprising categorizing the dialogues into different domains.
28. The method according to claim 27 further comprising ranking the categorized dialogues in each of the domains based on, at least in part, a frequency of utterance of each of the dialogues by the user over the period of time, such that the dialogues with a higher frequency of utterance are ranked higher as compared to the dialogues with a relatively lower frequency of utterance in the corresponding domain.
29. The method according to claim 28, wherein the identifying frequently uttered dialogues further comprises identifying a predefined number of dialogues with high rankings in each of the domains as the frequently uttered dialogues.
30. The method according to claim 27, wherein the domains include one or more of food, entertainment, sports, scheduling, sales inquiry, and automation commands.