CLAIMS

1. A computer-implemented method for providing an artificial intelligence (AI) -based digital assistant, the method being executed by one or more processors and comprising:

receiving, by an interaction manager, communication data from a device, the communication data comprising data input by a user of the device, the interaction manager coordinating interactions between the device and an action handler;

receiving, by the interaction manager, text data based on the communication data;

providing an intent set and an entity set based on processing the text data through an artificial intelligence service, the intent set comprising one or more intents indicated in the text data, the entity set comprising one or more entities indicated in the text data, the artificial intelligence service implementing one or more convolution neural networks (CNNs);

identifying, by the interaction manager, a set of actions based on one or more of the text data, the intent set, and the entity set, the set of actions comprising one or more actions to be executed by one or more third-party computer-implemented services;

transmitting, by the interaction manager, data representative of an action to the action handler, the action handler requesting execution of the action by a third-party computer implemented service;

receiving, by the interaction manager and from the action handler, a set of results comprising at least one result from the third-party computer-implemented service executing the action of the set of actions;

providing, by the interaction manager, result data comprising data describing the at least one result; and

transmitting, by the interaction manager, the result data to the device.

2. The method of claim 1, wherein the artificial intelligence service comprises an intent classification model using natural language processing (NLP) to provide the intent set.

3. The method of claim 2, wherein the NLP comprises word embedding.

4. The method of claim 1, wherein the artificial intelligence service comprises an entity extraction model using named entity recognition (NER) to provide the entity set.

5. The method of claim 1, further comprising determining that one or both of the intent set and the entity set is empty, and in response, transmitting at least one disambiguation question to the device.

6. The method of claim 1, further comprising determining that an expected entity is absent from the entity set based on an intent of the intent set, and in response, transmitting at least one disambiguation question to the device.

7. The method of claim 1, further comprising determining that the set of results includes a deficiency, and in response, transmitting at least one disambiguation question to the device.

8. The method of claim 1, wherein the communication data comprises audio data, and the result data comprises audio result data.

9. The method of claim 1, wherein the communication data comprises text data, and the result data comprises text result data.

10. The method of claim 1, wherein the result data comprises audio data that is provided by a voice response composition module based on text result data.

11. One or more non-transitory computer-readable storage media coupled to one or more processors and having instructions stored thereon which, when executed by the one or more processors, cause the one or more processors to perform operations for providing an artificial intelligence (AI) -based digital assistant, the operations comprising:

receiving, by an interaction manager, communication data from a device, the communication data comprising data input by a user of the device , the interaction manager coordinating interactions between the device and an action handler;

receiving, by the interaction manager, text data based on the communication data;

providing an intent set and an entity set based on processing the text data through an artificial intelligence service, the intent set comprising one or more intents indicated in the text data, the entity set comprising one or more entities indicated in the text data, the artificial intelligence service implementing one or more convolution neural networks (CNNs);

identifying, by the interaction manager, a set of actions based on one or more of the text data, the intent set, and the entity set, the set of actions comprising one or more actions to be executed by one or more third-party computer-implemented services;

transmitting, by the interaction manager, data representative of an action to the action handler, the action handler requesting execution of the action by a third-party computer implemented service;

receiving, by the interaction manager and from the action handler, a set of results comprising at least one result from the third-party computer-implemented service executing the action of the set of actions;

providing, by the interaction manager, result data comprising data describing the at least one result; and

transmitting, by the interaction manager, the result data to the device.

12. The computer-readable storage media of claim 11, wherein the artificial intelligence service comprises an intent classification model using natural language processing (NLP) to provide the intent set.

13. The computer-readable storage media of claim 12, wherein the NLP comprises word embedding.

14. The computer-readable storage media of claim 11, wherein the artificial intelligence service comprises an entity extraction model using named entity recognition (NER) to provide the entity set.

15. The computer-readable storage media of claim 11, wherein operations further comprise determining that one or both of the intent set and the entity set is empty, and in response, transmitting at least one disambiguation question to the device.

16. The computer-readable storage media of claim 11, wherein operations further comprise determining that an expected entity is absent from the entity set based on an intent of the intent set, and in response, transmitting at least one disambiguation question to the device.

17. The computer-readable storage media of claim 11, wherein operations further comprise determining that the set of results includes a deficiency, and in response, transmitting at least one disambiguation question to the device.

18. The computer-readable storage media of claim 11, wherein the communication data comprises audio data, and the result data comprises audio result data.

19. The computer-readable storage media of claim 11, wherein the communication data comprises text data, and the result data comprises text result data.

20. The computer-readable storage media of claim 11, wherein the result data comprises audio data that is provided by a voice response composition module based on text result data.

21. A system, comprising:

one or more processors; and

a computer-readable storage device coupled to the one or more processors and having instructions stored thereon which, when executed by the one or more processors, cause the one or more processors to perform operations for providing an artificial intelligence (AI) -based digital assistant, the operations comprising:

receiving, by an interaction manager, communication data from a device, the communication data comprising data input by a user of the device, the interaction manager coordinating interactions between the device and an action handler;

receiving, by the interaction manager, text data based on the communication data;

providing an intent set and an entity set based on processing the text data through an artificial intelligence service, the intent set comprising one or more intents indicated in the text data, the entity set comprising one or more entities indicated in

the text data, the artificial intelligence service implementing one or more convolution neural networks (CNNs);

identifying, by the interaction manager, a set of actions based on one or more of the text data, the intent set, and the entity set, the set of actions comprising one or more actions to be executed by one or more third-party computer-implemented services;

transmitting, by the interaction manager, data representative of an action to the action handler, the action handler requesting execution of the action by a third-party computer-implemented service;

receiving, by the interaction manager and from the action handler, a set of results comprising at least one result from the third-party computer-implemented service executing the action of the set of actions;

providing, by the interaction manager, result data comprising data describing the at least one result; and

transmitting, by the interaction manager, the result data to the device.

22. The system of claim 21, wherein the artificial intelligence service comprises an intent classification model using natural language processing (NLP) to provide the intent set.

23. The system of claim 22, wherein the NLP comprises word embedding.

24. The system of claim 21, wherein the artificial intelligence service comprises an entity extraction model using named entity recognition (NER) to provide the entity set.

25. The system of claim 21, wherein operations further comprise determining that one or both of the intent set and the entity set is empty, and in response, transmitting at least one disambiguation question to the device.

26. The system of claim 21, wherein operations further comprise determining that an expected entity is absent from the entity set based on an intent of the intent set, and in response, transmitting at least one disambiguation question to the device.

27. The system of claim 21, wherein operations further comprise determining that the set of results includes a deficiency, and in response, transmitting at least one disambiguation question to the device.

28. The system of claim 21, wherein the communication data comprises audio data, and the result data comprises audio result data.

29. The system of claim 21, wherein the communication data comprises text data, and the result data comprises text result data.

30. The system of claim 21, wherein the result data comprises audio data that is provided by a voice response composition module based on text result data.