What is claimed is :

1. A computing device comprising:

a processing unit;

memory;

one or more microphones; and

a speech-controlled digital personal assistant configured to perform operations comprising:

receiving speech input generated by a user via the one or more microphones;

performing speech recognition using the speech input to determine a spoken command, wherein the spoken command comprises a request to perform a task of a third-party application, and wherein the task is identified using a data structure that defines tasks of third-party applications invokable by spoken command, the third-party applications comprising tasks that are capable of being executed using user interfaces of the third-party applications and tasks that are capable of being headlessly executed without using the user interfaces of the third-party application, and the data structure defining how the task is capable of being executed by the digital personal assistant ;

using the data structure to select between headlessly executing the task of the third-party application and executing the task using the user interface of the third-party application ;

causing the third-party application to execute as a background process to headlessly execute the task when headless execution of the task of the third-party application is selected;

receiving a response from the third-party application indicating a state associated with the task; and

providing, by a user interface of the speech-controlled digital personal assistant, a response to the user based on the received state associated with the task so that the response comes from within a context of the user interface of the speech-controlled digital personal assistant without surfacing the user interface of the third-party application.

2. The computing device according to claim 1, wherein using the data structure to select between heedlessly executing the task of the third-party application and executing the task using the user interface of the third-party application comprises determining that the data structure defines the task as a background task.

3. The computing device according to claim 1, wherein using the data structure to select between heedlessly executing the task of the third-party application and executing the task using the user interface of the third-party application comprises determining that the user is using the computing device in a hands-free mode.

4. The computing device according to claim 1, the operations further comprising initiating a warm-up sequence of the third-party application while performing speech recognition and before completion of determining the spoken command.

5. The computing device according to claim 4, wherein the warm-up sequence includes allocating a portion of the memory, pre-fetching instructions, establishing a communication session, retrieving information from a database, starting a new execution thread, or raising an interrupt.

6. The computing device according to claim 1, wherein the data structure that defines tasks of third-party applications invokable by spoken command includes a mapping of states associated with the tasks to responses to the user.

7. The computing device according to claim 1, the operations further comprising:

starting a timer when the third-party application is caused to execute as a background process; and

terminating the background process if the timer expires.

8. The computing device according to claim 1, wherein the response from the third-party application indicates a confirmation state and the response to the user based on the confirmation state prompts the user to respond with a yes or no answer.

9. A method, implemented by a computing device comprising a microphone, the method comprising:

receiving, by a voice-controlled digital personal assistant, a digital voice input generated by a user, wherein the digital voice input is received via the microphone;

performing natural language processing using the digital voice input to determine a user voice command, wherein the user voice command comprises a request to perform a pre-defined function of a third-party voice-enabled application, and wherein the pre-defined function is identified using a data structure that defines functions supported by available third-party voice-enabled applications using voice input the third-party voice-enabled applications comprising pre-defined functions that are capable of being executed using user interfaces of the third-party voice-enabled applications and pre-defined functions that are capable of being heedlessly executed without using the user interface of the third-party voice-enabled applications, and the data structure defining how the pre-defined function is capable of being executed by the digital personal assistant ;

in response to determining the user voice command comprising the request to perform the pre-defined function of the third-party application, using the data structure to

select between heedlessly executing the pre-defined function of the third-party voice-enabled application and executing the pre-defined function using the user interface of the third-party voice-enabled application ;

causing the third-party voice-enabled application to execute the pre-defined function as a background process without a user interface of the third-party voice-enabled application appearing on a display of the computing device;

receiving a response from the third-party voice-enabled application indicating a state associated with the pre-defined function; and

providing, by a user interface of the voice-controlled digital personal assistant, a response to the user based on the received state associated with the pre-defined function so that the response comes from within a context of the user interface of the voice-controlled digital personal assistant without surfacing the user interface of the third-party voice-enabled application.

10. The method according to claim 9, further comprising initiating a warm-up sequence of the third-party voice-enabled application while performing natural language processing and before completion of determining the user voice command.

11. The method according to claim 10, wherein initiating the warm-up sequence comprises sending a speculative function to the third-party voice-enabled application.

12. The method according to claim 9, wherein the state associated with the pre-defined function is selected from a warm-up, initial, in-progress, confirmation, disambiguate, or final state.

13. The method according to claim 9, wherein the data structure is extensible via a command schema usable to associate requests from the user to pre-defined functions of one or more third-party voice-enabled applications.

14. The method according to claim 9, wherein the data structure comprises:

a first mapping of voice commands to functions supported by available third-party voice-enabled applications; and

a second mapping of states received from the functions to responses provided to the user from the voice-controlled digital personal assistant.

15. The method according to claim 9, wherein the responses provided to the user from the voice-controlled digital personal assistant are selected from the types : display text, text-to-speech, deep-link User Response Identifier (URI), web link URI, HyperText Markup Language (HTML), list template, physical address, or telephone number.

16. The method according to claim 9, wherein the third-party voice-enabled application is a remote web service.

17 The method according to claim 9, wherein performing natural language processing using the digital voice input to determine a user voice command comprises resolving contextual information so that the user voice command is context-free..

18. A computer-readable storage device storing computer-executable instructions for causing a computing device to perform operations for completing a task of a voice-enabled application within the context of a voice-controlled digital personal assistant, the operations comprising:

receiving, by the voice-controlled digital personal assistant, a digital voice input generated by a user, wherein the digital voice input is received via a microphone;

performing natural language processing using the digital voice input to determine a user voice command, wherein the user voice command comprises a request to perform the task of the voice-enabled application, and wherein the task is identified using an extensible data structure that maps user voice commands to tasks of voice-enabled applications, the voice-enabled applications comprising tasks that are capable of being executed using user interfaces of the voice-enabled applications and tasks that are capable of being heedlessly executed without using the user interfaces of the voice-enabled applications, and the data structure defining how the task is capable of being executed by the digital personal assistant ;

in response to determining the user voice command, using the data structure to select between executing the task of the voice-enabled application as a foreground task or a background task;

when the task is selected to be executed as a background task, causing the voice-enabled application to execute the task as a background task and within a context of the voice-controlled digital personal assistant without a user interface of the voice-enabled application surfacing;

receiving a response from the voice-enabled application indicating a state associated with the task; and

providing a response to the user based on the received state associated with the task, the response provided within the context of the voice-controlled digital personal assistant without a user interface of the voice-enabled application surfacing when it is determined that the task is a background task.

19.. The computer-readable storage device according to claim 18, the operations further comprising initiating a warm-up sequence of the voice-enabled application while performing natural language processing and before completion of determining the user voice command.