

Task 2

Google's Assistant and NLP

Google Assistant uses Natural Language Processing (NLP) and below is a general step-by-step overview:

Audio capture: the system continuously listens for the wake-up words “OK Google” or “Hey Google” so the user can interact with the virtual assistant. The system once activated by one of those words begins listening and capturing the audio.

Speech Recognition: Automatic Speech Recognition (ASR) technology is used to process the recorded audio. ASR transcribes the speech by converting the spoken words into written text.

Intent Recognition: Following the transcription of the speech, NLP algorithms examine the text to ascertain the user's intent or the action they wish to take. This requires comprehending the query's structure, significance, and context.

Language Understanding: NLP models interpret user queries by gathering pertinent data and identifying key entities, like places, times, or particular requests. In this step, the text is parsed and mapped to a structured representation the system can use.

Action Fulfilment: The system chooses the right course of action based on the intent as it has been interpreted and the information it has extracted. It might entail looking up information on the internet, running a particular command, or interacting with other services or applications.

Response Generation: After the requested action has been chosen, the system either generates a response or carries it out. Depending on the device and user interface, the response may take the form of spoken words, displayed text, or a combination of the two.

It's important to remember that, to maintain privacy and minimise latency, wake-up word detection and audio capture typically take place on the device itself. However, because they call for substantial computational resources and access to expansive language models, the speech recognition, language understanding, and response generation stages frequently involve cloud-based processing.

For "Hey Google," Google's NLP technologies are based on machine learning methods such as deep neural networks. Large amounts of labelled data are used to train these models, which are then refined over time by incorporating a variety of linguistic patterns and user interactions.

In general, the "Hey Google" system, which is powered by NLP, combines speech recognition, language understanding, and action fulfilment to offer a conversational and interactive experience with Google Assistant and allow users to interact with their devices through voice commands.

References

<https://blog.google/products/assistant/ask-a-techspert-assistant-questions/>

<https://www.ibm.com/topics/natural-language-processing>

<https://aws.amazon.com/what-is/nlp/>

<https://www.encora.com/insights/natural-language-processing-and-machine-learning>