building a conversation agent

I love the idea of talking to my phone for more than just making calls. Tools like Alexa, Google Assistant and others allow you to create voice interactions, either one off or repeating. In this post I'll use Google Assistant to create a voice search interface for this blog as an Google Assistant agent. We'll explore Voice User Interfaces, How to integrate the blog's search functionality with Assistant and how well it works.

Voice User Interface

Tools like Siri, Google Assistant and Alexa provide 2 way voice conversation between the app and the user through custom actions or programs to accomplish single or groups of tasks.

This is different than using [Speech Recognition](#) and [Speech Synthesis API](#) to create communication between users and the apps

Google Assistant, Actions on Google, and Dialog Flow

I will concentrate in Google Assistant, the Google competitor to Cortana (from Microsoft) and Alexa (from Amazon) for these examples.

Before we start we'll define some of the technologies and products that we'll use in this project and

Assistant
The application that we use actions with, either with voice of the keyboard

Actions On Google
Actions on Google is the platform that lets you create software (actions) to extend the functionality of the Google Assistant

Action
The actual applications that provide conversational interfaces to products or

services.

The entry point into an interaction that you build for the Assistant. Users can request your Action by typing or speaking to the Assistant.

Dialog Flow

The tool used to build actions for Google Assistant

You can code the functions entirely on Dialog Flow using built-in intents or you can create web hooks to connect to third party services

Cloud Functions

Google Cloud Functions is a serverless execution environment for building and connecting cloud services. Your Cloud Function is triggered when an event being watched is fired and it executes in a fully managed environment

In this project I've chosen to work with Cloud Functions for Firebase rather than host my own server to host them

Building the action

The high-level concept for the action is as follows:

1. The user requests the action
2. The action replies with an acknowledgment and an initial request for words to search
3. The user speaks or types the term they want to search
4. The action triggers the cloud function and the function provides a response including the search results
5. The action replies with the results from the function

To Publish or not to publish

Conclusion

Links and resources

- Designing for Interaction Modes
- Designing Voice User Interfaces — Cathy Pearl — O'Reilly

- [Conversational Design](#) — Erika Hall — A Book Apart
- [Conversations with Robots: Voice, Smart Agents & the Case for Structured Content](#)