

Assistent personal per l'EPSEVG basat en Alexa

Miriam Martínez Izquierdo

Escola Politècnica Superior d'Enginyeria de Vilanova i la Geltrú

Resum

L'ús dels assistents per veu s'ha tornat una rutina diària per moltes persones. Aquests permeten una interacció per veu amb els diferents dispositius que el tenen integrat.

Donada aquesta situació, es planteja el desenvolupament d'una skill de l'assistent per veu Alexa, de la companyia Amazon. Una skill és una funcionalitat extra per l'assistent, equivalent a una aplicació per mòbils. Aquesta skill permetrà als usuaris obtenir informació relacionada amb l'activitat acadèmica de l'EPSEVG.

Finalment, el resultat aconseguit en el treball és una skill que proporciona diferents dades públiques sobre l'EPSEVG i per altra banda, una documentació detallada sobre el funcionament d'Alexa i les skills, que permetrà la publicació i manteniment d'aquesta en un futur.

1. Introducció

Els assistents per veu faciliten la interacció entre els usuaris i els diferents dispositius que l'integren. Permeten la realització de tasques d'una manera senzilla i fàcil, similar a la interacció humana. Per aquest motiu, l'objectiu del treball és donar a conèixer aquesta tecnologia. La intenció del treball és oferir un servei d'atenció per a la comunitat de l'Escola i a possibles visitants amb informació d'interès sobre la mateixa.

Per a l'assoliment d'aquest objectiu, es realitzaran les següents fases. Un estudi de l'arquitectura de la plataforma Amazon que fa possible la creació de skills, el disseny de la skill plantejada contemplant tant el seu funcionament com la definició dels diferents casos d'ús, l'estudi de les eines a utilitzar i el seu desenvolupament. Posteriorment, es realitzarà la fase de validació, on es comprovarà el correcte funcionament de la skill i es definiran els passos a seguir pel seu manteniment.

2. Skills d'Alexa

Les skills d'Alexa són les diferents funcionalitats que pot oferir l'assistent per veu Alexa. Les skills que es poden crear a la plataforma que ofereix Amazon permeten ampliar les funcionalitats bàsiques que ofereix Alexa inicialment.

Existeixen diferents tipus de skills, en la skill desenvolupada en aquest treball s'ha escollit la custom skill. El que permet aquest tipus de skill és una total personalització de la interacció entre l'usuari i Alexa.

Les skills estan formades per dues parts, la part de front end que gestionaria la interacció amb l'usuari, gràcies a un model d'interacció. I la part de back end que és l'encarragada del tractament de les peticions realitzades per l'usuari i el posterior retorn de la informació sol·licitada.

El model d'interacció per una skill és l'estructura que determina les sol·licituds a les que pot donar resposta i tenir una descripció de les diferents paraules que pot dir un usuari per invocar aquestes sol·licituds. Aquesta estructura queda representada en un arxiu JSON, aquest inclou totes les possibilitats d'interacció amb l'usuari disponibles per la skill.

Per la creació del model d'interacció s'han de definir les següents parts:

- Nom d'invocació: nom amb el que l'usuari podrà invocar a la skill.
- Intents: les sol·licituds a les que la skill podrà donar resposta. Aquests estan formats per:
 - Sample utterances: frases probables que l'usuari pot dir per cridar un intent.
 - Slots: arguments variables dins d'un sample utterance per a poder conèixer la informació específica que busca l'usuari.

La creació d'aquest model ens permet tenir un control sobre la interacció entre l'usuari i la skill.

3. Funcionament

El model d'interacció és l'encarregat de descriure a quines peticions dona resposta Alexa i el servei del back end és l'encarregat de donar la resposta a Alexa perquè la sintetitzi.

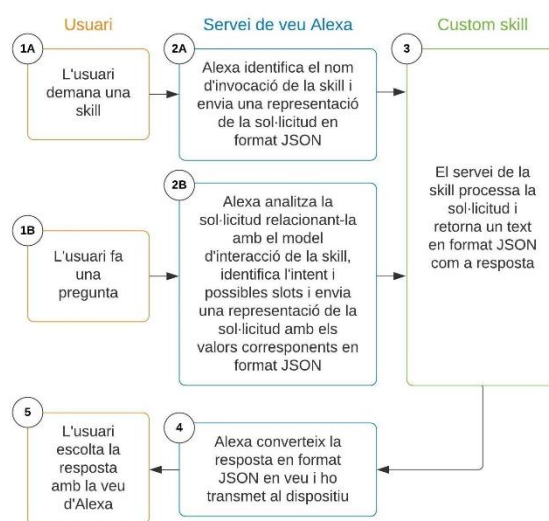


Fig. 1. Diagrama de flux entre l'usuari i una custom skill

4. Disseny

Per a facilitar la posterior implementació, s'han de definir una sèrie de casos d'ús per la skill:

1. Informació professors: la responsabilitat d'aquest és retornar les dades que sol·licita un usuari sobre un determinat professor de l'Escola. Aquestes dades poden ser el número de telèfon, despatx, horari de consultes i el departament al que pertany.
2. Horari de consulta professors: la responsabilitat d'aquest cas d'ús és retornar les dades sol·licitades sobre l'horari de consultes del curs actual d'un professor determinat. Aquesta informació serà el dia de la setmana, l'aula, hora d'inici i fi de les hores de consulta.
3. Informació departaments: la responsabilitat d'aquest és retornar les dades demanades sobre un departament de l'Escola. Aquestes dades poden ser el codi del departament, el nom o el nom del responsable del departament a l'Escola.
4. Informació aules: la responsabilitat d'aquest cas d'ús és retornar l'edifici i la planta on es troba una determinada aula de l'Escola.
5. Informació setmana actual: la responsabilitat és retornar el número de setmana i ordre del curs actual.
6. Horari assignatures: la responsabilitat és retornar la informació disponible sobre els horaris d'una assignatura del curs i quadrimestre actuals. La informació pot ser el dia de la setmana, hora d'inici i fi, aula i grup.
7. Informació exàmens: la responsabilitat és retornar les dades del tipus d'examen demanat per una assignatura. Pot ser un examen parcial, final o de reavaluació. Retornarà la data, hora i aula.
8. Informació lectures TFG: la responsabilitat és la de retornar les dades demanades per l'usuari, en aquest cas pot ser la informació sobre la lectura d'un estudiant o sobre les lectures previstes per un professor en el dia de realitzar la consulta.

Una vegada definits aquests cas d'ús tenim una idea més clara de les diferents peticions a les que podrà donar resposta la skill complint amb les responsabilitats definides.

5. Implementació

Les custom skills estan formades per un model d'interacció, front end i un servei, back end. L'arquitectura de la skill creada, *Asistente EPSEVG*, és la que mostra la Fig. 2.

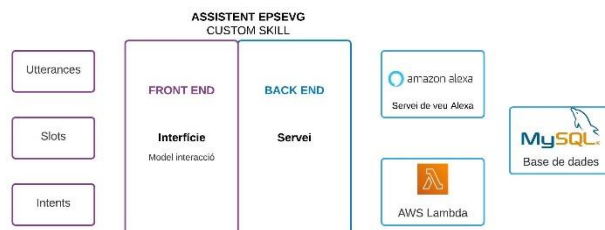


Fig. 2. Arquitectura Asistente EPSEVG

Com es pot veure a la Fig. 2, la part de front end és el model d'interacció i el back end està format pel servei de veu d'Alexa[1], una funció Lambda executada a l'Amazon Web Services[2] i una còpia de la base de dades MySQL de l'Escola.

La creació del model d'interacció s'ha pogut realitzar d'una manera més organitzada i senzilla gràcies al disseny previ dels casos d'ús, cada intent creat correspon a un cas d'ús dels definits. Els sample utterances són les frases possibles que podria dir un usuari en base a la informació que pot proporcionar cada intent. Els slots és la informació variable que pot tenir, com per exemple els noms dels professors pels que demana informació.

El back end consisteix en una funció Lambda amb codi escrit en Javascript. En aquest codi es troba un fitxer encarregat de controlar els intents els que crida un usuari mitjançant un SDK que ofereix Alexa Skills Kit. Cada funció d'aquest fitxer amb el codi corresponent a la crida d'un intent que cridarà a les funcions d'un altre fitxer que conté les funcions que es connecten i consulten a la base de dades, i retornarà la informació demanada. Aquesta resposta amb la informació demanada es retorna en un arxiu JSON com mostra la Fig.3 on el que sintetitzarà Alexa és la variable *outputsSpeech*.

```
"body": {
  "version": "1.0",
  "response": {
    "outputSpeech": {
      "type": "SSML",
      "ssml": "<speack>El profesor Daniel Guasch Murillo partici"
    },
    "reprompt": {
      "outputSpeech": {
        "type": "SSML",
        "ssml": "<speack>¿Puedo ayudar en algo más?</speack>"
      }
    }
  }
},
```

Fig. 3. Exemple de resposta JSON generada pel back end

6. Validació

Una vegada implementada la skill, es realitza una validació final. S'utilitzen dues formes per a poder comprovar el seu correcte funcionament.

1. Simulador d'Alexa[3]: mitjançant aquest simulador podem veure de manera visual i per veu el flux d'interacció de l'usuari i la skill. A més de veure els arxius JSON, l'arxiu on descriu la petició que realitza l'usuari i els diferents paràmetres i l'altre la resposta generada pel back end.

2. Altaveu intel·ligent Echo Dot: mitjançant aquest dispositiu podem fer una comprovació del funcionament interactuant directament amb ell. Analitzant les respostes que sintetitza Alexa i comprovant la correcta experiència d'usuari[4].

7. Conclusions

El treball realitzat conté una estructura on cada apartat assoleix una de les fases plantejades com a objectius inicialment.

En primer lloc, s'ha aconseguit la realització d'un estudi sobre els assistents per veu, més concretament Alexa i la seva plataforma per a desenvolupar skills. Gràcies a la realització d'aquesta part s'ha pogut aprofundir en l'arquitectura d'aquesta plataforma, concretament en l'arquitectura i funcionament d'una custom skill.

En segon lloc, s'ha aconseguit la realització d'un estudi sobre les diferents eines que ofereix Amazon per al desenvolupament d'una skill. També es va realitzar una primera aproximació a les futures eines per al desenvolupament del back end de la skill.

En tercer lloc, la valoració de quines dades mostrades a la pàgina web de l'Escola eren més rellevants en aquest context d'Alexa. Una vegada decidit, realitzar una especificació amb els diferents casos d'ús i els esdeveniments de sistema que podria tractar cadascun. És a dir, el disseny de la skill.

En quart lloc, la decisió final de les eines a utilitzar per a la implementació i la posterior implementació de la skill. Una vegada comprovada la possible utilització d'una còpia de la base de dades de l'Escola per a l'obtenció de les dades, es va començar a implementar. Aquesta implementació ha estat completa satisfactòriament.

Per últim, la validació de la skill. Per tant, la comprovació del correcte funcionament del desenvolupat durant el projecte. I a la vegada, la comprovació de la correcta aplicació dels coneixements sobre les eines utilitzades, l'arquitectura i funcionament de la plataforma Alexa.

En conclusió, es pot dir que s'han assolit els objectius plantejats inicialment de manera satisfactòria. Però evidentment el cicle de la skill no s'ha tancat, per tant seguidament s'exposaran les possibles futures tasques i millores a realitzar.

Primerament, la skill no ha estat publicada a la botiga de skills d'Amazon. Per tant, el següent pas seria publicar-la, pujar-la a producció. Com que aquesta skill seria un producte oficial de l'Escola i s'ha d'elaborar una política de privacitat com a requisit, no s'ha realitzat aquest pas.

En segon lloc, el model d'interacció és una part fonamental a la skill. Donat que el servei de veu d'Alexa interactua millor o pitjor en gran part en funció d'aquest model creat de zero en el cas de la custom skill. Per tant, seria molt bona opció seguir treballant en ell per a poder millorar-lo.

Per últim, a les primeres fases d'estudi sobre el funcionament de les skills i diferents eines disponibles es va plantejar la idea de poder utilitzar una API que tractés les dades de la base de dades depenent de la necessitat i no obtenir-les directament. Aquest canvi seria una millora molt important en termes de seguretat, donat que, una API seria un extra de

seguretat. Una API ofereix més seguretat que un accés directe a la base de dades.

8. Agraïments

M'agradaria agrair al director del treball, Rafael Vidal Ferré i al codirector, Daniel Guasch Murillo, per haver donat els feedbacks constructius corresponents durant l'elaboració del treball i per transmetre una motivació constant. Això últim, tret característic dels dos i motiu d'haver contactat inicialment per la realització del meu treball final de grau.

Per altra banda, un agraïment especial a Víctor Agramunt, de serveis TIC de l'Escola. Sense ell no hagués estat possible la obtenció de les dades de la base de dades de l'Escola, aspecte essencial en l'elaboració d'aquest treball. Destacar també la seva gran implicació i entusiasme durant el desenvolupament del treball, la seva disposició a realitzar reunions per a resoldre dubtes sobre l'estructura de les dades entre d'altres.

Finalment, agrair a tots els professors i companys que he arribat a tenir durant tota l'etapa, perquè part de tot el que he après al final és producte de cadascun d'ells. Tant la base teòrica i pràctica adquirida, com tots els punts de vista i consells rebuts.

Referències

- [1] Funcionament servei de veu d'Alexa, AVS. Documentació de la pàgina web d'Amazon Developer. [Consulta: 12/10/2020]
<https://developer.amazon.com/en-US/alexa/devices/alexa-built-in/development-resources/sdk>
- [2] AWS Lambda. Documentació de la pàgina web d'Amazon Web Services. [Consulta: 28/12/2020]
https://docs.aws.amazon.com/es_es/lambda/latest/dg/welcome.html
- [3] Testing d'una skill. Pàgina web d'Amazon Developer. [Consulta: 02/01/2021]
<https://developer.amazon.com/en-US/docs/alexa/devconsole/test-your-skill.html>
- [4] Vídeo propi de la validació de la skill. Pàgina web Google Drive. [Consulta: 28/01/2021]
https://drive.google.com/file/d/1aXUFtk8aFCSaY5vm2_Hi1X4s6aSOonv5/view?usp=sharing

Bibliografia

Documentació oficial Amazon Developer
<https://developer.amazon.com/es-ES/docs/alexa/ask-overviews/what-is-the-alexa-skills-kit.html>