



UNIVERSITATEA POLITEHNICA TIMIȘOARA

Facultatea de Electronică, Telecomunicații și
Tehnologii Informaționale Timișoara



PROIECT DE DIPLOMĂ

POSIBILITĂȚI DE UTILIZARE A ASISTENȚILOR DE VORBIRE PE SISTEMUL RASPBERRY PI

Conducător științific:

Profesor dr. ing. Mischie Septimiu

Absolvent:

Popescu Dragoș

Timișoara
2018

Sinteza lucrării

Lucrarea de diplomă este dezvoltată în cadrul Departamentului de Electronică Aplicată al Universității Politehnica din Timișoara. Lucrarea va constitui rezultatul efortului depus pe parcursul mai multor luni, în perioada octombrie 2017 - iunie 2018.

Această lucrare se încadrează într-o tematică de interes pentru domeniul recunoașterii vocale și al asistenților vocali din momentul actual. Scopul lucrării este studierea și portarea asistenților vocali Amazon Alexa și Google Assistant pe Raspberry Pi, în vederea controlului unor dispozitive hardware, cu ajutorul comenzilor vocale.

Capitolul 1 face o introducere a lucrării și a sistemul Raspberry Pi .

Capitolul 2 face o introducere despre asistenții vocali disponibili pe piață.

Capitolul 3 prezintă pașii necesari instalării Amazon Alexa pe sistemul Raspberry Pi.

Capitolul 4 prezintă pașii necesari instalării Google Assistant pe sistemul Raspberry Pi.

Capitolul 5 prezintă shield-ul Voice Hat folosit la Raspberry Pi pentru a facilita folosirea Google Assistant pe Raspberry Pi.

Capitolul 6 prezintă implementarea algoritmului pentru controlul unor dispozitive hardware cu ajutorul comenzilor vocale și rezultatele obținute cu acesta.

Capitolul 7 prezintă concluziile de rigoare, extrase în urma rezultatelor obținute prin implementarea algoritmului, dar și referințele bibliografice.

Cuprins

SINTEZA LUCRĂRII.....	2
CAP. 1: INTRODUCERE	4
1.1 Raspberry Pi prezentare generala.....	5
1.2 Raspberry Pi 2 Model B vs. Raspberry Pi 3 Model B	6
1.3 Hardware.....	8
1.4 Porturile USB.....	9
1.5 Portul HDMI	9
1.6 Portul Ethernet	9
1.7 Software	9
CAP.2: Asistenți vocali	10
2.1 Google Assistant:	13
2.2 Siri:	14
2.3 Bixby:	16
2.4 Amazon Alexa:	18
CAP.3 Instalarea Amazon Alexa pe sistemul Raspberry Pi.	19
CAP.4 Instalarea Google Assistant pe Raspberry Pi.	26
CAP.5 Google AIY Voice Kit	39
CAP.6 Implementarea algoritmilor pentru controlul unor dispozitive hardware	44
CAP.7 Concluzii	51

CAP. 1: INTRODUCERE:

Am ales această lucrare deoarece mi-a permis studiul unui domeniu relativ nou și de actualitate, care nu a fost îndelung studiat, cel puțin în țara noastră și anume asistenții vocali și recunoașterea vocală .

Am dorit să îmbin recunoașterea vocală cu controlul unor echipamente hardware(În speță un kit de robot).dar și să evidențiez posibilitatea utilizării lor pe o platformă Raspberry Pi datorită costului redus pe care îl are un Raspberry Pi față de orice device cu asistenți vocali disponibil pe piață în acest moment .



Sursă : aiyprojects.withgoogle.com/voice-v1/

Figura 1 : Google Aiy project



Sursă: beebom.com

Figura 2 : Diverse-uri cu asistenți vocali

Domeniul roboticii a fost încă de la începuturi și este și în ziua de astăzi unul dintre cele mai interesante direcții de specializare pentru un electronist, dar totodată este și unul dintre cele mai complexe deoarece se află la intersecția mai multor științe de sine stătătoare precum mecanica, electronica, știința calculatoarelor, chimia prin știința materialelor, fizica și lista se lungeste pe zi ce trece! Condițiile care au dus la apariția roboților sunt economice, de mediu, iar tehnologia din ziua de azi împinge robotica și spre domeniul mircotehnologiei și nanotehnologiei [11].

Divertismentul este și el un domeniu ce dă naștere unor roboți, existând ligi de lupte între roboți incluzând faimoșii sumo-boți, ligi de fotbal pentru roboți iar în următorii ani se va lansa și competiția de mașini electrice autonome de curse, în mărime naturală, Roborace[10].



Sursa: roborace.com

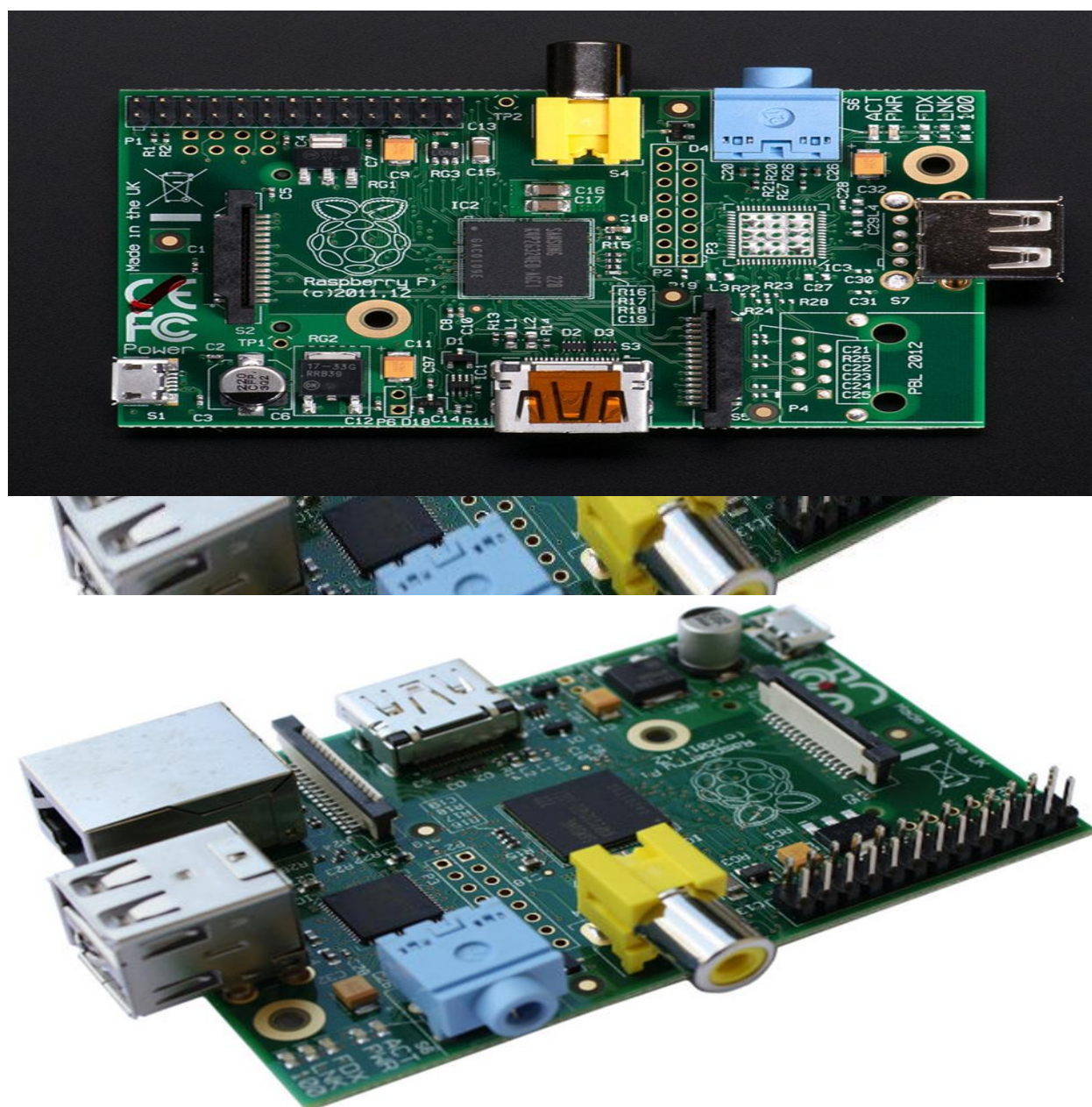
Figura 3: Mașina oficială Roborace

1.1. Raspberry Pi prezentare generala

Raspberry Pi este un computer de mărimea unui card de credit, dezvoltat în Marea Britanie de către Raspberry Pi Foundation, cu scopul de a promova învățarea noțiunilor de bază din domeniul informaticii în școli.

Caracteristicile sale și platforma open source, bazată pe Linux, îl fac versatil, potrivit pentru aplicații obișnuite, precum home entertainment, navigare internet, vizionare multi-media, light gaming, dar este cel mai folosit în aplicații de dezvoltare, precum automatizari, robotica, etc.

Raspberry Pi a fost creat în 2012, primele modele lansate fiind: Model A, cu un processor de 700MHz, 256 Mb memorie RAM, un singur port USB, 26 pini hardware, acesta ne reprezentând un port Ethernet și Model B, cu un processor de 700MHz, 512 Mb memorie RAM, 2 porturi USB, 26 de pini hardware și port Ethernet.



Sursă : vesalia.de

Figura 5 : Raspberry Pi model B

De-a lungul timpului, au fost lansate mai multe modele îmbunătățite de Raspberry: Raspberry Pi Model A+, Raspberry Pi 2 model B, Raspberry Pi 3 model B.

Raspberry permite conectarea la un monitor, tastatură și mouse, putând să fie programat astfel utilizând diverse limbaje de programare cum ar fi: Python, Scratch sau C/C++. Acesta este capabil să se comporte precum un computer normal, având avantajul unui preț de achiziție mult mai scăzut față de un computer, dar și avantajul portabilității.

În proiectul meu am folosit Raspberry Pi 3 Model B.

1.2 Raspberry Pi 2 Model B vs. Raspberry Pi 3 Model B :



Sursă:adafruit.com

Figura 6: Raspberry Pi 2 model B vs Raspberry Pi 3 model B



Sursă:google.ro

Figura 7 : Raspberry Pi 1 model B vs Raspberry Pi 2 model B vs Raspberry Pi 3 model B

Raspberry Pi 2 Model B are următoarele caracteristici:

- 1 GB memorie RAM
- procesor quadcore de 900 MHZ
- 4 porturi USB
- 1 port Ethernet
- 1 port HDMI
- mufa jack 3.5 mm
- 40 pini GPIO
- Sistem de operare Raspbian Linux 8

Raspberry Pi 3 Model B a fost lansat ulterior și prezintă câteva îmbunătățiri ale caracteristicilor față de versiunea 2:

- 1 GB memorie RAM
- procesor quadcore de 1.2 GHZ
- 4 porturi USB
- 1 port Ethernet
- 1 port HDMI
- mufa jack 3.5 mm
- 40 pini GPIO
- Sistem de operare Raspbian Linux 9
- Permite conectivitate Bluetooth și WI-FI.

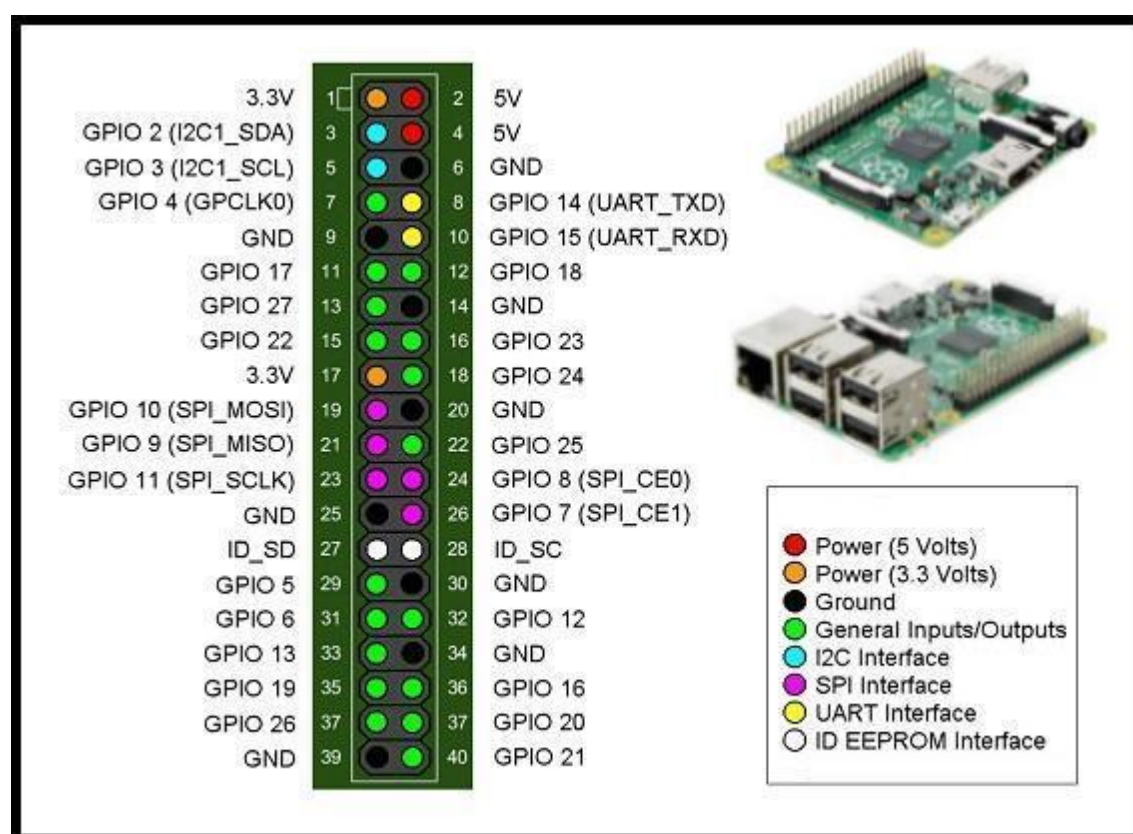
Se observă din caracteristicile de mai sus că Raspberry Pi 3 are o viteză a procesorului mai mare și permite conectivitate Bluetooth și WI-FI.

1.3 Hardware

1.3.1 Pini GPIO

Aceștia sunt pini de intrare/iesire (GPIO-General Purpose Input/Output), care pot fi controlați cu ajutorul software-ului. Putem conecta la ei diferite periferice (ex. Led-uri, Senzori, Motoare DC-DC), pe care le putem controla prin aplicații software. Pinii GPIO pot să fie folosiți și ca pini de comunicare prin diferite interfețe: SPI,I2C,UART. De asemenea, ei permit accesul la tensiuni de 3.3V și 5V și la GND.

Configurația pinilor pentru Raspberry PI 3 este prezentată în Fig.2.3.1 de mai jos:



Sursă:google.ro

Figura 8 : Pinii Raspberry Pi 2 și 3 , primii 26 fiind prezenți si la Raspberry Pi 1

1.4 Porturile USB

Raspberry Pi este echipat cu porturi USB care permit conexiunea prin USB a diferitor periferice. În acest proiect am folosit aceste porturi pentru a conecta tastatura, mouse-ul și microfonul.

1.5 Portul HDMI

Raspberry Pi prezintă un port HDMI prin care este posibilă conectarea la ecranul unui PC, TV.

Deoarece monitorul folosit nu prezintă o ieșire HDMI, a fost necesară folosirea unui adaptor de la HDMI la DVI, corespunzător cu ieșirea monitorului .

1.6 Portul Ethernet

Conectarea la rețeaua locală a device-ului Raspberry Pi a fost făcută prin intermediul WI-FI-ului disponibil pe Raspberry Pi 3, dar se putea utiliza și cablul de internet, care trebuia conectat la portul Ethernet de pe Raspberry.

Prin intermediul WI-FI-ului, am realizat conexiunea între Router și Raspberry, atribuindu-se o adresă IP pentru Raspberry PI, de forma 192.168.1.yyy.

Pentru a afla adresa IP, în terminalul Raspberry trebuie să tastăm următoarea comandă:

- Ifconfig

1.7 Software

Raspberry Pi are ca sistem de operare o distribuție de Linux, denumită Raspbian.

Linux reprezintă o familie de sisteme de operare de tip Unix, care folosesc Nucleul Linux (kernel). Acesta poate fi instalat pe o varietate largă de Hardware: telefoane mobile, tablete, console video.

Raspbian este un sistem de operare gratuit, bazat pe Debian și optimizat pentru hardware-ul Raspberry Pi, dezvoltat de o echipă de programatori, pasionați de proiectul Raspberry Pi al Raspberry Foundation , care vine cu peste 40000 de pachete preinstalate.

Pentru a putea vedea exact sistemul de operare care rulează pe Raspberry-ul nostru, în terminal trebuie să tastăm următoarea comandă:

-cat/etc/os-release

CAP.2: Asistenți vocali

Contextul apariției:

Vorbirea este un mijloc versatil de comunicare. Ea este compusă din : mesaje lingvistice (de exemplu, Mesaj și limbă), vorbitor (de exemplu, caracteristicile emoționale și fiziologice ale aparatul vocal) și mediul înconjurător (de exemplu, unde a fost produs și transmis discursul). Chiar dacă astfel de informații sunt codificate într-o formă complexă, oamenii le pot descifra relativ ușor pe majoritatea.

Această abilitate umană a inspirat cercetătorii să elaboreze sisteme care să imite o astfel de abilitate. De la foneticieni la ingineri, cercetătorii au lucrat pe mai multe fronturi pentru a decoda majoritatea informațiilor din semnalele acustice. Unele fronturi includ sarcini cum ar fi identificarea vorbitorului prin recunoașterea vocii, detectarea limbii vorbite, transcrierea discursului, traducerea discursului și înțelegerea vorbirii.

Recunoașterea vocală este folosită pe scară largă pentru diverse aplicații în diverse gadget-uri și sisteme electronice. Aplicația principală este de a susține căutarea în baza de date cu ajutorul serviciilor IVRS (Sistem integrat de răspuns vocal). Aplicația IVRS îi ajută pe utilizatori să caute informațiile fie prin selectarea tactilă a meniurilor, fie prin intermediul comenzilor vocale.

Recunoașterea automată a vorbirii (ASR) a fost în centrul atenției pentru mulți cercetători de mai multe decenii. În acest tip de recunoaștere, mesajul lingvistic este informația de interes. Aplicațiile de recunoaștere a vorbirii variază de la dictarea unui text la generarea de subtitrări în timp real pentru o emisiune de televiziune. În ciuda capacității umane, cercetătorii au aflat că extragerea informațiilor din discurs nu este un proces simplu. Variabilitatea în vorbire datorită limbajului lingvistic, factori fiziologici și de mediu îi provoacă pe cercetători să extragă fiabil informațiile relevante din semnalul de vorbire. În ciuda tuturor provocărilor, cercetătorii au făcut progrese semnificative în tehnologie, astfel încât este posibil să se dezvolte aplicații cu funcție de vorbire.

Problema recunoașterii vorbirii poate fi descrisă ca o funcție, care definește o descriere de la dovezile acustice la o singură sau la o succesiune de cuvinte. Fie $X = (x_1, x_2, x_3, \dots, x_t)$ reprezintă dovezile acustice, care sunt generate în timp (indicate de indicele t) dintr-un semnal de vorbire dat și aparțin setului complet de secvențe acustice, χ . Fie $W = (w_1, w_2, w_3, \dots, w_n)$ o secvență de cuvinte n , fiecare aparținând unui set fix și cunoscut de cuvinte posibile, ω . Există două cadre pentru a descrie funcția de recunoaștere a vorbirii: șablon și statistică.

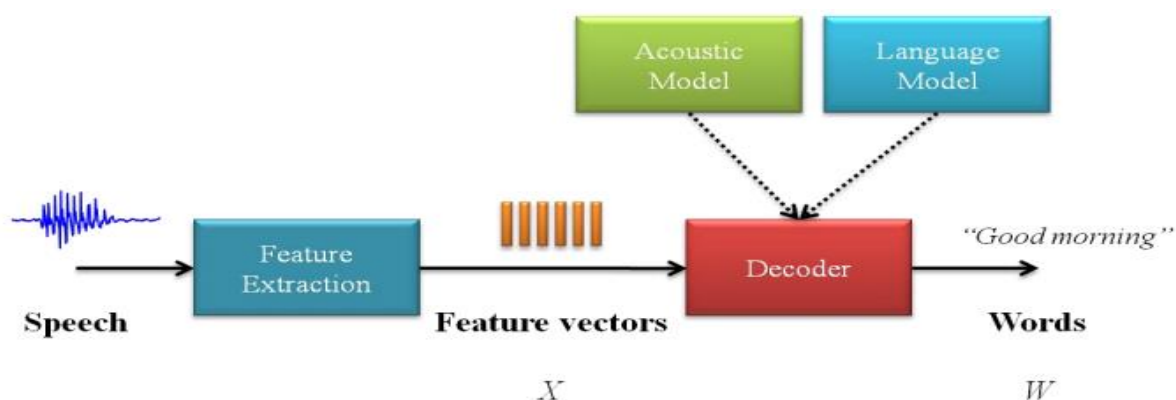
În cadrul statistic, recunoașterea selectează secvența de cuvinte care este mai probabil să fie produsă, având în vedere dovezile acustice observate. Cu toate acestea, deoarece $P(W|X)$ sunt dificil de modelat direct, regula Bayes ne permite să rescriem o astfel de probabilitate

$$P(W|X) = \frac{P(W)P(X|W)}{P(X)}$$

unde $P(W)$ este probabilitatea ca să se rostească secvența de cuvinte W , $P(X|W)$ este probabilitatea de a observa probele acustice X atunci când vorbitorul pronunță W , și $P(X)$ este probabilitatea ca dovezile acustice X să fie observate. Termenul $P(X)$ poate fi abandonată deoarece este o constantă. Apoi, recunoașterea ar trebui să selecteze secvența de cuvinte W care maximizează produsul $P(W) * P(X|W)$, adică:

$$\tilde{W} = \underset{W \in \omega}{\operatorname{argmax}} P(W)P(X|W). \quad (1)$$

Cele mai reușite sisteme de recunoaștere a vorbirii se bazează pe cadrul statistic descrise în secțiunea anterioară. Ecuația (1) stabilește componentele unei recunoașteri vocale. Probabilitatea anterioară $P(W)$ este determinată de un model de limbă, iar probabilitatea $P(X|W)$ este determinată de un set de modele acustice, iar procesul de căutarea peste toate secvențele posibile ale cuvintelor W care maximizează produsul este realizat de decodor. Figura prezintă componentele principale ale unui sistem ASR.



Sursă: rfwireless-world.com

Figura 9: Schema bloc a unui sistem de recunoaștere vocală automat

Figura descrie diagrama simplă a sistemului de recunoaștere a vorbirii. Așa cum este prezentat, utilizatorul vorbește, din eșantioanele prelevate se extrage anumiți vectori de dimensiuni specifice care sunt stocați și decodați în text pe baza modelelor acustice, limbajului și a genului.

- Recunoașterea vorbirii detectează mai întâi și captează cuvinte vorbite.
- Le transformă în reprezentare digitală după eliminarea zgomotului. Acest lucru se face folosind algoritmi DSP(Digital Signal Processor sau procesor de semnal digital (numeric)).
- Împărțirea sunetelor în bucăți mai mici dacă sunt eșantioane foarte mari.
- Acum, discursul este atribuit unor fenomene particulare bazate pe înregistrări în bazele de date și pe baza probabilităților. Recunoașterea automată a vorbirii este, în esență, o conversie independentă a limbii vorbite în text citibil în timp real. Următoarele sunt avantajele ale sistemului automat de recunoaștere a vorbirii:
- Reducerea costurilor datorită automatizării acestui sistem.
- Capabilitatea de a căuta text
- Pregătire pentru utilizare de către oamenii cu probleme auditive și vizuale.

Începând din 2017, capacitățile și utilizarea asistenților virtuali se extind rapid, introducând noi produse pe piață. Un sondaj online din mai 2017 a constatat că cele mai utilizate pe scară largă în SUA sunt Siri (34%), Asistent Google (19%), Amazon Alexa (6%) și Microsoft Cortana (4%). Apple și Google sunt instalate de utilizatori pe smartphone-uri. Microsoft are o bază mare de instalări pe computere personale bazate pe Windows, smartphone-uri și difuzoare inteligente. Alexa are o bază mare de instalare mare pentru difuzoare inteligente.

Primul instrument activat pentru recunoașterea vocală digitală a fost IBM Shoebox, prezentat publicului larg în timpul târgului mondial din Seattle, în 1962, după lansarea sa inițială pe piață în 1961. Acest computer timpuriu s-a dezvoltat cu aproape 20 de ani înainte de introducerea primului computer personal IBM în 1981 a reușit să recunoască 16 cuvinte vorbite și cifrele 0 la 9. Următoarea etapă în dezvoltarea tehnologiei de recunoaștere a vocii a fost realizată în anii 1970 la Universitatea Carnegie Mellon din Pittsburgh, Pennsylvania, cu sprijinul substanțial al Departamentului de Apărare a Statelor Unite ale Americii și agenția sa DARPA. Instrumentul lor "Harpy" a stăpânit aproximativ 1000 de cuvinte, aproximativ vocabularul unui copil de trei ani. Zece ani mai târziu, același grup de oameni de știință a dezvoltat un sistem care ar putea analiza nu numai cuvintele individuale, ci și secvențe întregi de cuvinte activate de un model Markov ascuns. Astfel, cei mai vechi asistenți virtuali, care aplicau software-ul de recunoaștere a vorbirii, au fost programele automate de dictare și de asistență medicală. În anii 1990 tehnologia de recunoaștere a vorbirii digitale a devenit o caracteristică a calculatorului personal cu Microsoft, IBM, Philips și Lernout & Hauspie care luptau pentru clienți. Mult mai târziu, lansarea pe piață a primului smartphone IBM Simon în 1994 a pus bazele asistenților virtuali inteligenți așa cum îi cunoaștem astăzi. Primul asistent virtual digital instalat pe un smartphone a fost Siri, care a fost introdus ca o caracteristică a iPhone 4S pe 4 octombrie 2011. Apple Inc. a dezvoltat Siri ca urmare a achiziției din 2010 a companiei Siri Inc., o parte a SRI International, un institut de cercetare finanțat de DARPA și Departamentul Apărării al Statelor Unite

Iată și principalii „asistenți vocali inteligenți” pe care îi putem folosi la ora actuală:

- Pe Android: Google Assistant
- Pe iOS: Siri
- Pe Samsung Galaxy S8: Bixby
- Amazon Alexa

2.1 Google Assistant:

Asistentul Google este un asistent virtual dezvoltat de Google, care este disponibil în primul rând pe dispozitive mobile și smart home. Spre deosebire de Google Now, Asistentul Google se poate angaja în conversații în ambele sensuri.

Asistentul a debutat inițial în mai 2016, ca parte a aplicației de mesagerie Google Allo și Google Home. După o perioadă de exclusivitate pe smartphone-urile Pixel și Pixel XL, a început să fie instalată pe alte dispozitive Android în februarie 2017, inclusiv smartphone-uri terțe și Android Wear, și a fost lansată ca aplicație independentă în sistemul de operare iOS în luna mai. Alături de anunțul unui kit de dezvoltare software în aprilie 2017, Asistentul a fost și este extins în continuare pentru a sprijini o mare varietate de dispozitive, inclusiv mașini și aparate electrocasnice inteligente. Funcționalitatea Asistentului poate fi îmbunătățită și de dezvoltatorii terți.

Utilizatorii interacționează în primul rând cu Asistentul Google prin voce naturală, deși este acceptată și tastatura. În aceeași natură și în același mod ca Google Now, Asistentul poate căuta pe Internet, programa evenimente și alarme, poate ajusta setările hardware pe dispozitivul utilizatorului și poate afișa informații din contul Google al utilizatorului. Google a anunțat, de asemenea, că Asistentul va putea să identifice obiecte și să adune informații vizuale prin camera aparatului și să sprijine achiziționarea de produse și trimiterea de bani, precum și identificarea melodiilor.

La CES 2018, au fost anunțate primele display-uri inteligente cu funcții auxiliare (difuzoare inteligente cu ecrane video), cu o lansare planificată pentru mijlocul anului 2018. [8]



Sursă: google.ro

Figura 10 : Sigla Google Assitant

2.2 Siri:

Siri (pronunțat / sɪəri / SEER-ee) este un asistentă virtuală parte a sistemelor de operare iOS, watchOS, macOS și tvOS ale Apple Inc. Asistentul utilizează interogări vocale și o interfață utilizator în limbaj natural pentru a răspunde la întrebări, a face recomandări și a efectua acțiuni prin delegarea cererilor la un set de servicii pe Internet. Software-ul se adaptează utilizatorilor la utilizările lingvistice individuale, căutări și preferințe, cu utilizare continuă. Rezultatele returnate sunt individualizate.

Siri are la bază un proiect inițial dezvoltat de către SRI International Center Intelligence Artificial. Motorul său de recunoaștere vocală a fost furnizat de Nuance Communications, iar tehnologia Siri utilizează tehnologii avansate de deep learning pentru a funcționa. Actorii ei originali americani, britanici și australieni au înregistrat vocea lor în jurul anului 2005, fără să știe despre eventuala utilizare a înregistrărilor în Siri. Asistentul vocal a fost lansat ca o aplicație pentru iOS în februarie 2010 și a fost achiziționat de Apple două luni mai târziu. Siri a fost apoi integrat în iPhone 4S la lansarea sa în octombrie 2011. În acel moment, aplicația separată a fost, de asemenea, eliminată din iOS App Store. Siri a devenit o parte integrantă a produselor Apple, după ce a fost adaptată în alte dispozitive hardware de-a lungul anilor, inclusiv modele iPhone mai noi, iPad, iPod Touch, Mac, AirPods, Apple TV și HomePod.

Siri acceptă o gamă largă de comenzi pentru utilizatori, inclusiv efectuarea de acțiuni telefonice, verificarea informațiilor de bază, planificarea evenimentelor și a mementourilor, gestionarea setărilor dispozitivelor, căutarea pe Internet, navigarea pe zone, găsirea informațiilor despre divertisment și posibilitatea de a se angaja cu aplicațiile integrate iOS. Odată cu lansarea iOS 10 în 2016, Apple a deschis accesul limitat la terțe părți la Siri, inclusiv aplicațiile de mesagerie terță parte, precum și plățile, aplicațiile de partajare a apelurilor și apelurile pe Internet. Odată cu lansarea iOS 11, Apple a actualizat vocea lui Siri pentru mai multe voci clare, umană, traducerea de limbi străine și acțiuni terțe suplimentare.

Descărcarea originală a lui Siri pe iPhone 4S în 2011 a primit recenzii mixte. Acesta a primit o laudă pentru recunoașterea vocală și cunoașterea contextuală a informațiilor despre utilizatori, inclusiv pentru întâlnirile din calendar, însă a fost criticat pentru necesitatea unor comenzi rigide de utilizator și lipsa de flexibilitate.

A fost, de asemenea, criticat pentru lipsa de informații despre anumite locuri din apropiere și pentru incapacitatea sa de a înțelege anumite accente britanice. În 2016 și 2017, o serie de rapoarte

mass-media au indicat că Siri este împotriva inovației, în special împotriva noilor asistenți vocali concurenți din alte companii de tehnologie. Rapoartele au vizat un set limitat de caracteristici ale lui Siri, recunoașterea vocii "proaste" și integrarea serviciilor nedezvoltate ca provocând probleme Apple în domeniul inteligenței artificiale și serviciilor bazate pe cloud; baza pentru reclamațiile care se pare că se datorează dezvoltării greoaie, cauzată de prioritizarea de către Apple a confidențialității utilizatorilor și a luptelor de putere a executivului în cadrul companiei. [7]



Sursă: apple.com

Figura 11: Sigla Siri

2.3 Bixby:

Bixby este un asistent virtual dezvoltat de Samsung Electronics.

În 20 martie 2017, Samsung a anunțat asistentul digitală cu voce numită "Bixby". Bixby a fost introdus alături de Samsung Galaxy S8 și S8 + în cadrul evenimentului Samsung Galaxy Unpacked 2017, care a avut loc pe 29 martie 2017. Samsung a dezvăluit oficial Bixby cu o săptămână înainte de lansare, dar a făcut prima apariție doar în timpul evenimentului. Bixby poate fi, de asemenea, utilizat pe dispozitivele Galaxy mai vechi care rulează Android Nougat.

Bixby reprezintă un restart major pentru S Voice, aplicația de asistență vocală Samsung introdusă în 2012 cu Galaxy S III.

În mai 2017, Samsung a anunțat că Bixby va veni pe linia sa de frigidere Family Hub 2.0, devenind primul produs care nu include mobilul care să includă asistentul virtual.

În octombrie 2017, Samsung a anunțat lansarea lui Bixby 2.0 în cadrul conferinței sale anuale de dezvoltare din San Francisco. Noua versiune va fi lansată pe linia de produse lansate de companie, inclusiv smartphone-uri, televizoare și frigidere. În plus, terților li se va permite să dezvolte aplicații pentru Bixby utilizând Kitul de dezvoltare Samsung.

Bixby Voice

Cu Bixby Voice, utilizatorul poate declanșa Bixby apelând-o sau apăsând lung butonul Bixby situat sub tastatura de volum. Cu ceva timp înainte de lansarea telefonului, butonul Bixby putea să fie reprogramat și putea fi setat să deschidă alte aplicații sau asistenți, cum ar fi Google Assistant. Aproape de lansarea telefonului, această abilitate a fost eliminată cu o actualizare de firmware, dar poate fi remapată folosind aplicații terțe.

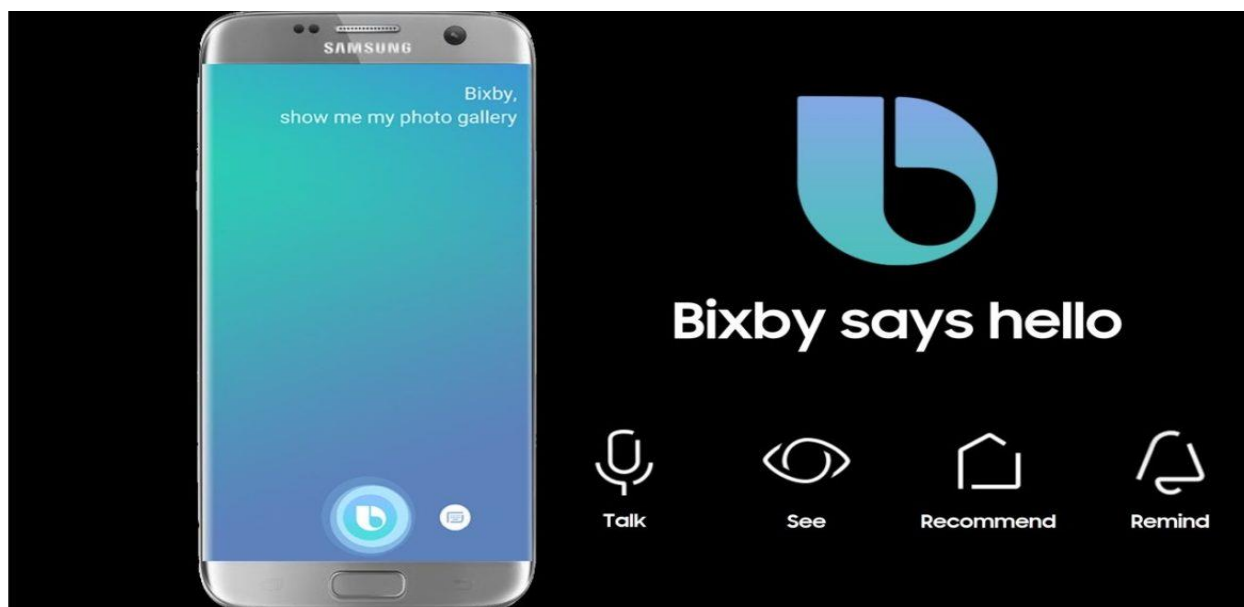
Bixby Vision

Bixby Vision este încorporată în aplicația camerei fiind în esență o cameră de realitate augmentată, care poate identifica obiecte în timp real, le poate căuta pe diverse servicii și să ofere utilizatorului posibilitatea de a le cumpăra, dacă acestea sunt disponibile. Bixby este de asemenea capabil să traducă text, să citească coduri QR și să recunoască repere.

Bixby Home

Bixby Home poate fi găsit prin deplasarea spre dreapta pe ecranul de întâmpinare sau prin apăsarea scurtă a butonului Bixby. Este o listă verticală a informațiilor pe care Bixby le poate interacționa, de exemplu, cu vremea, cu activitățile de fitness și cu butoanele pentru controlul gadget-urilor inteligente de acasă.

Bixby acceptă mai multe limbi, inclusiv engleza, coreeană și chineză. De asemenea, acceptă căutarea contextuală și căutarea vizuală. [6]



Sursă:samsung.com
Figura 12:Sigla Bixby

2.4 Amazon Alexa:

Alexa este un asistent virtual dezvoltat de Amazon, utilizat pentru prima oară în Amazon Echo și difuzoarele inteligente Amazon Echo Dot dezvoltate de Amazon Lab126. Este capabil să interacționeze cu vocea, să redea muzică, să facă liste de lucru, să stabilească alarme, să facă streaming podcast-uri, să cânte audiobook-uri și să ofere informații despre vreme, trafic, sport și alte informații în timp real, cum ar fi știrile. Alexa poate controla, de asemenea, mai multe dispozitive inteligente, utilizându-se ca sistem de automatizare a locuinței. Utilizatorii pot extinde capacitățile Alexa instalând "abilități" (funcționalități suplimentare dezvoltate de furnizori terți, în alte setări mai des numite aplicații cum ar fi programe meteo și caracteristici audio).

Majoritatea dispozitivelor cu Alexa permit utilizatorilor să activeze dispozitivul folosind un cuvânt de avertizare (cum ar fi Echo); alte dispozitive (cum ar fi aplicația mobilă Amazon pe iOS sau Android) solicită utilizatorului să apese un buton pentru a activa modul de ascultare al lui Alexa. În prezent, interacțiunea și comunicarea cu Alexa sunt disponibile numai în engleză, germană și japoneză. În noiembrie 2017, Alexa a devenit disponibilă pe piața canadiană numai în limba engleză.

În septembrie 2017, Amazon avea mai mult de 5.000 de angajați care lucrau pe Alexa și pe produse conexe.

În noiembrie 2014, Amazon a anunțat apariția Alexa alături de Echo. Alexa a fost inspirată din sistemul de voce și de conversație la bordul companiei Starship Enterprise în seriale și filme de știință science fiction, începând cu Star Trek: Seria originală și Star Trek: The Next Generation.

Numele Alexa a fost ales din cauza faptului că are o asemănare puternică cu cu X și, prin urmare, ar putea fi recunoscut cu o mai mare precizie. Numele, de asemenea, amintește de Biblioteca din Alexandria, care este de asemenea folosită de Amazon Alexa Internet din același motiv. În iunie 2015, Amazon a anunțat Alexa Fund, un program care va investi în companii care fac abilități și tehnologii de control vocal. Fondurile în valoare de 100 de milioane de dolari au fost investite în companii precum Ecobee, Chef de Orange, Scout Alarm, Garageio, Toymail, MARA și Mojio.

În ianuarie 2017, prima conferință Alexa a avut loc la Nashville, Tennessee, o adunare independentă a comunității mondiale de dezvoltatori și entuziaști Alexa. Urmărirea a fost anunțată, pentru a fi notată de către liderul original al produsului Amazon Alexa / Connected Home, Ahmed Bouzid.

În cadrul conferinței Amazon Web Services Re: Invent din Las Vegas, Amazon a anunțat Alexa ca fiind disponibil pentru afaceri și capacitatea dezvoltatorilor de aplicații de a-și plăti add-on-uri pentru abilitățile lor.

În mai 2018, Amazon a anunțat că Alexa va fi inclusă în toate cele 35.000 de case noi Lennar Corporation construite în acest an. [9]



Sursă: amazon.com

Figura 13: Sigla Amazon Alexa

CAP.3 Instalarea Amazon Alexa pe sistemul Raspberry Pi.

Pentru a folosi Amazon Alexa pe Raspberry Pi avem nevoie de următoarele componente:

- Un Raspberry Pi 3
 - Un cablu MicroUSB pentru alimentare
 - Un card de memorie MicroSD de minim 8GB
 - Un microfon pe USB
 - Boxe sau căști conectate pe jack-ul de 3.5mm
 - O tastatură și un mouse pentru a putea controla sistemul Raspberry Pi
- În continuare voi prezenta pașii necesari instalării de Amazon Alexa.

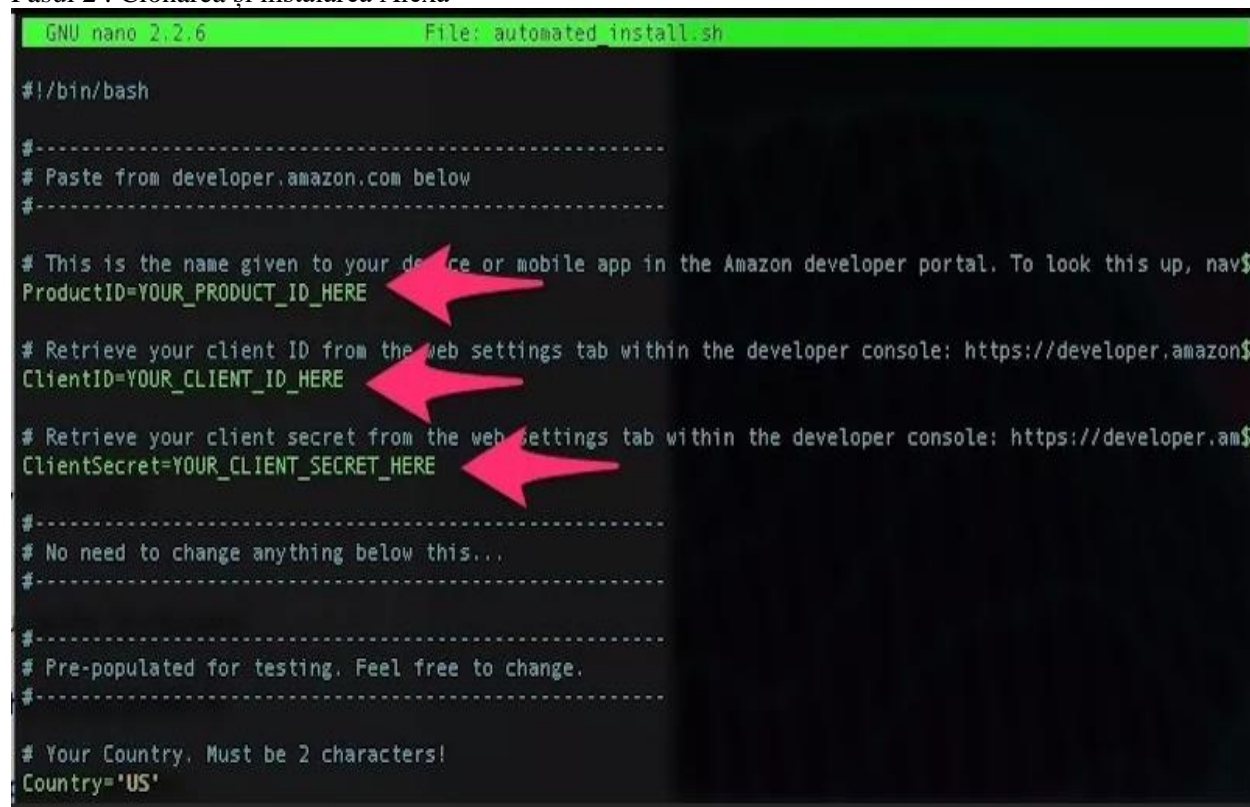
Pasul 1: Ne creăm cont pentru dezvoltator pe site-ul Amazon si executăm următoarele:

- Ne logăm în contul creat pentru dezvoltator pe site-ul Amazon
- Facem click pe tab-ul Alexa
- Facem click pe înregistrarea unui tip de produs și alegem dispozitivul
- Vom denumi tipul produsului și numele cu care acesta va apărea mai departe
- Apăsăm pe next
- În ecranul cu securitatea profilului facem click pe creează un profil nou
- Sub tabul general, următorul lângă Numele Profilului de Securitate, vom da numele profilului nostru. Vom face același lucru și pentru descriere, după care apăsăm next
- Ne vom nota Product ID, Client ID și Client Secret pe care site-ul ni le generează
- Apăsăm tabul de setări web , apoi click pe butonul de edit lângă bara de dropdown a profilului
- Lângă Allowed Origins, facem click pe “Add Another” și tastăm: <https://localhost:3000>
- Lângă Allowed Return URLs, facem click pe “Add Another” și tastăm : <https://localhost:3000/authresponse>
- Tabul de Device Details este următorul . Nu prea contează ce introducem aici, alegem o categorie, scriem o descriere, alegem un interval de timp și introducem 0 pe formularul de lângă dispozitivele pe care intenționăm să le utilizăm, după care facem click pe următorul.
- În final, putem alege să adăugăm Amazon Music, dar vom selecta “No” pentru ca acesta nu funcționează pe Raspberry Pi, după care facem click pe save.

The screenshot shows the Amazon Developer Console interface for creating a new device type. The top navigation bar includes the Amazon logo, the text 'DEVELOPER CONSOLE', the user name 'THORIN KLOSOWSKI - REPUBLIC OF THORONIA', and links for 'SIGN OUT' and 'ENGLISH'. Below the navigation bar is a menu with tabs: DASHBOARD, APPS & SERVICES, ALEXA (selected), REPORTING, SUPPORT, DOCUMENTATION, and SETTINGS. The main content area is titled 'Create a new Device Type' and includes a 'Back to the list' link. On the right, there are links for 'Getting started', 'AVS Agreement', 'AVS Program Requirements', and 'AVS Content Requirements'. The left sidebar shows a list of steps: 'Device Type Info' (checked), 'Security Profile' (checked), 'Device Details' (unchecked), and 'Amazon Music' (unchecked). The 'Security Profile' section is active, showing a description: 'You need a security profile to identify your device. Your security profile credentials - client ID and client secret - allow your device to securely identify itself to the Alexa Voice Service. If you are building a website, click here to [Learn More](#). If you are building an Android or iOS app, click here to [Learn More](#).' Below this, there is a 'Security Profile' dropdown menu set to 'PiLexa' with a 'Cancel' button. The 'General' tab is selected, showing fields for 'Allowed Origins' and 'Allowed Return URLs'. The 'Allowed Origins' field contains 'https://localhost:3000' and has an 'Add Another' button. The 'Allowed Return URLs' field contains 'https://localhost:3000/authresponse' and also has an 'Add Another' button. At the bottom left, a message states 'Successfully created security profile.' and at the bottom right, there is a 'Next' button.

Figura 14: Print screen realizat din pagina de creare a contului de Amazon

Pasul 2 : Clonarea și instalarea Alexa



```
GNU nano 2.2.6      File: automated_install.sh

#!/bin/bash

#-----
# Paste from developer.amazon.com below
#-----

# This is the name given to your device or mobile app in the Amazon developer portal. To look this up, nav$
ProductID=YOUR_PRODUCT_ID_HERE

# Retrieve your client ID from the web settings tab within the developer console: https://developer.amazon$
ClientID=YOUR_CLIENT_ID_HERE

# Retrieve your client secret from the web settings tab within the developer console: https://developer.am$
ClientSecret=YOUR_CLIENT_SECRET_HERE

#-----
# No need to change anything below this...
#-----

#-----
# Pre-populated for testing. Feel free to change.
#-----

# Your Country. Must be 2 characters!
Country='US'
```

Figura 15: Captură din scriptul automated_install.sh pentru introducerea id-urilor de utilizator

Conectați-vă totul la Raspberry Pi și porniți-l. Va trebui să fiți în interfața grafică de utilizator pentru că, în cele din urmă, utilizați un browser web pentru a vă autentifica dispozitivul.

- Deschidem Terminalul de la Raspberry și tastăm: `cd Desktop` după care apăsăm enter.
- Apoi tastăm `git clone https://github.com/alexa/alexa-avs-sample-app.git` după care apăsăm enter.
- Imediat ce clonarea s-a încheiat, tastăm: `cd ~/Desktop/alexa-avs-sample-app` și apăsăm enter.
- Tastăm `nano automated_install.sh` și apăsăm enter.
- Această comandă va deschide editorul de text. Aici, va trebui să introduceți ID-ul Produsului(Product Id), ClientID(Client ID) și Client Secret pe care le-am notat la pasul unu. Utilizați tastele săgeți pentru a naviga la fiecare intrare. Introduceți fiecare detaliu după semnul =, după cum se menționează în imaginea de mai sus. Când ați terminat, atingeți CTRL + X pentru a salva și a ieși
- După ce am salvat, ne întoarcem înapoi la linia de comandă din terminal. Vom rula script-ul de instalare prin tastarea comenzii `cd ~/Desktop/alexa-avs-sample-app` urmat de apăsarea tastei enter, după care tastăm `. automated_install.sh` și apăsăm din nou tasta enter.
- Când vi se solicită, apăsați pe Y pentru diferite întrebări și răspundeți după cum doriți pentru restul. Acest lucru va configura Raspberry Pi-ul dumneavoastră. și instalarea câtorva software-uri suplimentare. Acest lucru poate dura până la 30 de minute, așa că lăsați-l să facă acest lucru.

Odată ce termină, este timpul să începeți serviciul Alexa.

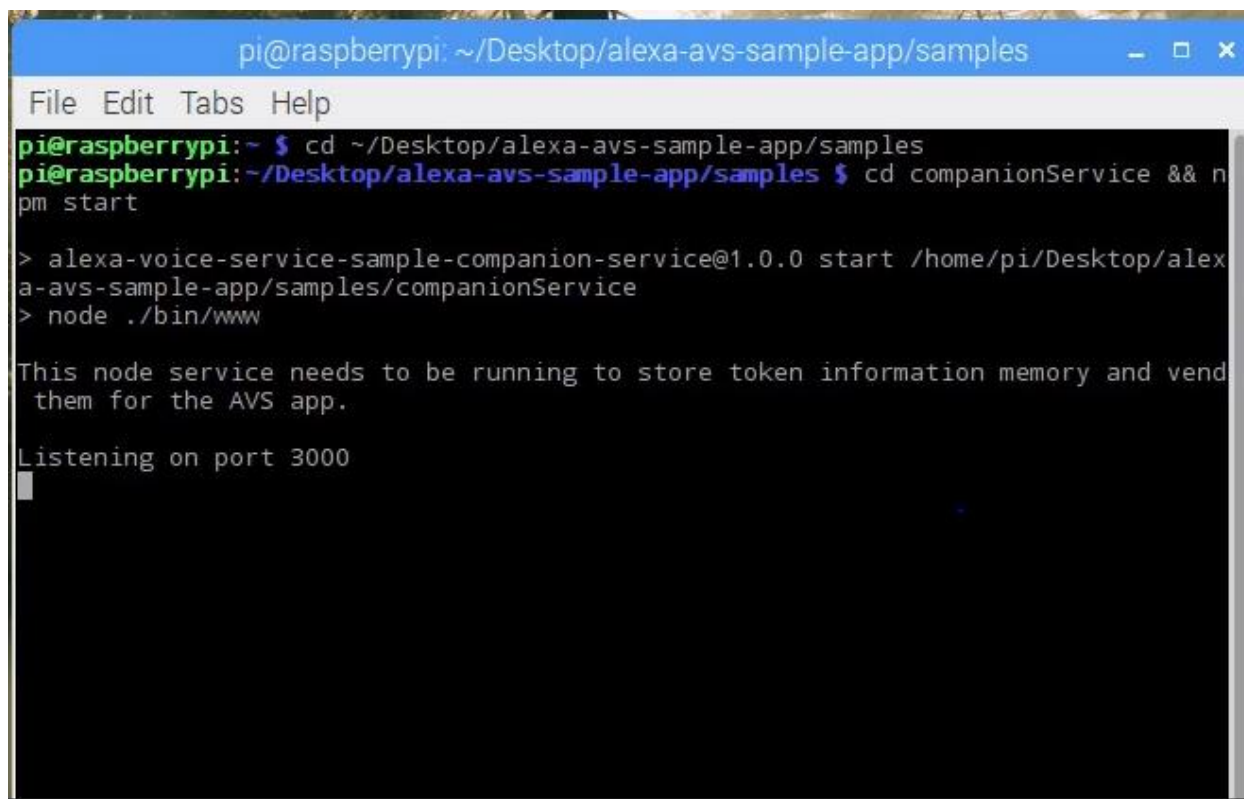
Pasul 3: Rularea Alexa Web Service

Apoi, vom executa trei seturi de comenzi simultan în trei ferestre Terminal diferite. Vom crea o nouă fereastră Terminal pentru fiecare din următorii pași. Nu vom închide ferestrele deschise când vom deschide una nouă! Vom avea nevoie să facem pașii, trei (acesta), patru și cinci de fiecare dată când repornim Raspberry Pi.

Primul pe care îl veți începe este Alexa Web Service:

- 1.Vom scrie `cd ~/Desktop/alexa-avs-sample-app/samples` si vom apăsa enter.
- 2.Vom scrie `cd companionService && npm start` si vom apăsa enter.

Acești pași pornesc Companion Service si deschid un port de comunicație cu Amazon. Lăsați fereastra deschisă!



```
pi@raspberrypi: ~/Desktop/alexa-avs-sample-app/samples
File Edit Tabs Help
pi@raspberrypi:~ $ cd ~/Desktop/alexa-avs-sample-app/samples
pi@raspberrypi:~/Desktop/alexa-avs-sample-app/samples $ cd companionService && npm start

> alexa-voice-service-sample-companion-service@1.0.0 start /home/pi/Desktop/alexa-avs-sample-app/samples/companionService
> node ./bin/www

This node service needs to be running to store token information memory and vend them for the AVS app.

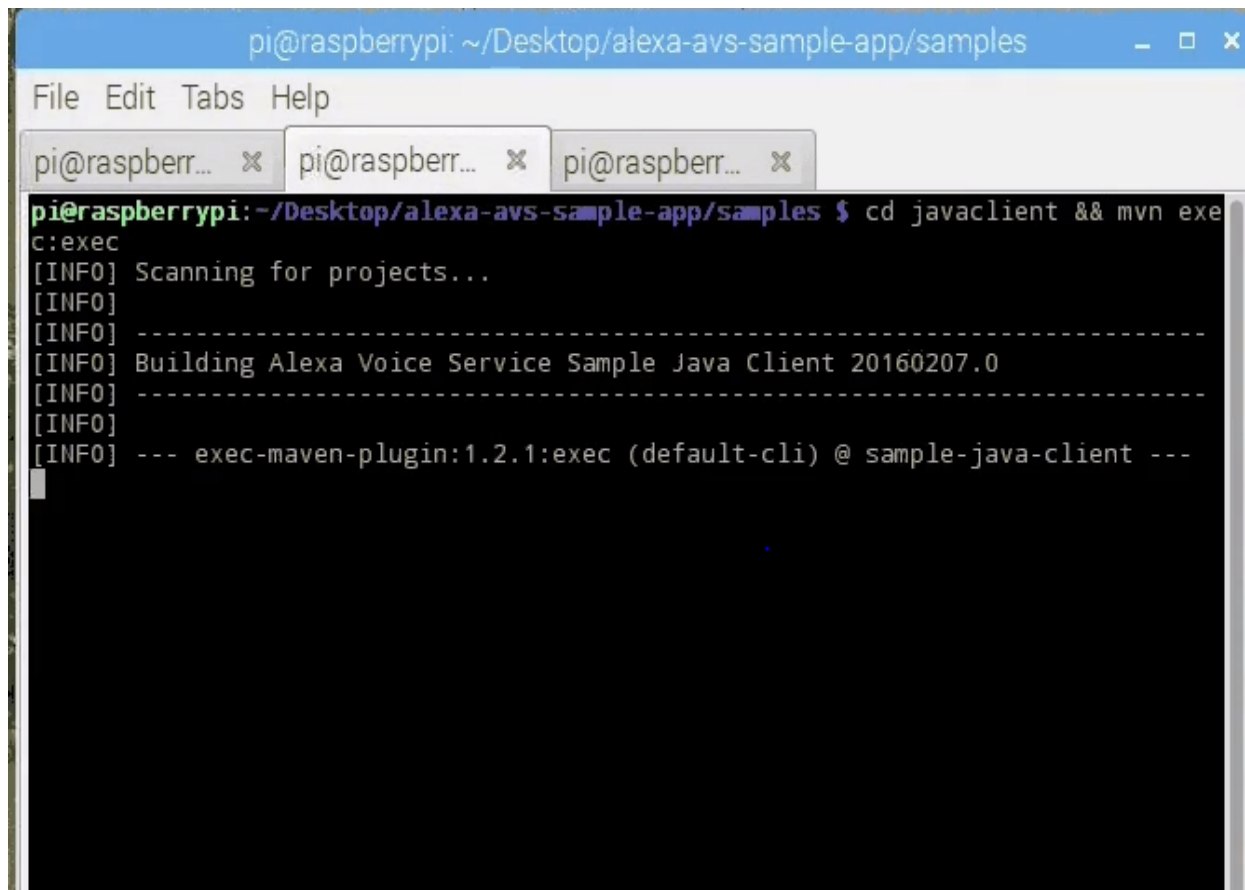
Listening on port 3000
```

Figura 16: Lansarea scriptului pentru pornirea companionService și deschiderea portului de comunicație cu Amazon

Pasul 4: Rulăm Sample App si ne confirmăm contul.

Deschidem a doua fereastră Terminal (File > New Window). Acest pas următor rulează o aplicație Java și lansează un browser web care înregistrează fișierele Echo cu ajutorul serviciului web Alexa(Alexa Web Service).

- În nouă fereastra terminal deschisă tastăm `cd ~/Desktop/alexa-avs-sample-app/samples` și apăsăm enter.
- Tastăm `cd javaclient && mvn exec:exec` și apăsăm enter.
- Va apărea o fereastră care vă cere să vă autentificați dispozitivul. Facem clic pe Da. Aceasta deschide o fereastră de browser. Un al doilea pop-up va apărea în aplicația Java cerând să faceți clic pe Ok. Nu faceți clic pe acesta încă.
- Ne logam cu contul de Amazon creat anterior in fereastra browser-ului.
- Vom vedea ecranul pentru autentificare pentru dispozitivul nostru . Facem click pe OK .Browser-ul va afișa acum “device tokens ready”.
- Acum putem sa facem pe click pe OK in pop-up-ul din aplicația Java. Acum, Raspberry Pi are jetoanele necesare pentru a comunica cu serverul Amazon. Lăsați această fereastră Terminal deschisă.



```
pi@raspberrypi: ~/Desktop/alexa-avs-sample-app/samples
File Edit Tabs Help
pi@raspberr... x pi@raspberr... x pi@raspberr... x
pi@raspberrypi:~/Desktop/alexa-avs-sample-app/samples $ cd javaclient && mvn exe
c:exec
[INFO] Scanning for projects...
[INFO]
[INFO] -----
[INFO] Building Alexa Voice Service Sample Java Client 20160207.0
[INFO] -----
[INFO]
[INFO] --- exec-maven-plugin:1.2.1:exec (default-cli) @ sample-java-client ---
```

Figura 17: Lansarea scriptului care va deschide fereastra de autentificare spre Amazon Alexa

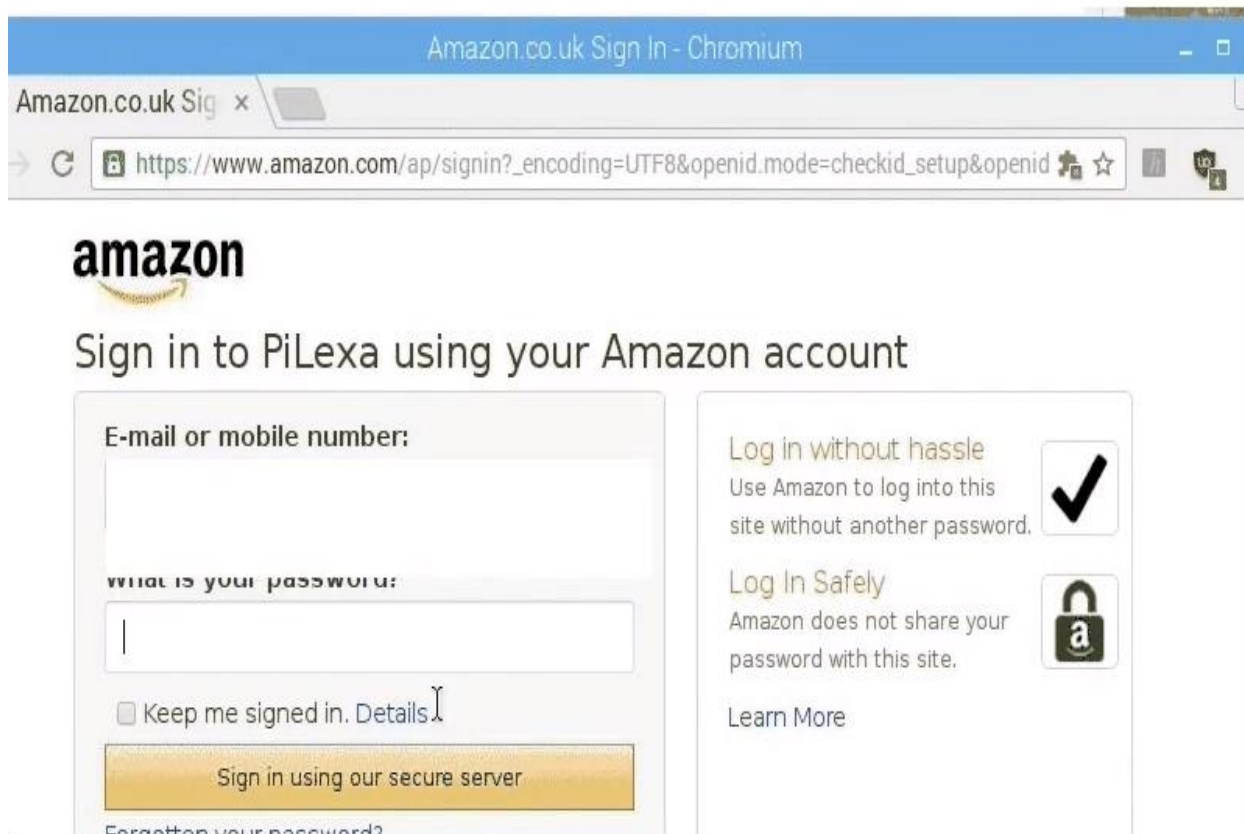


Figura 18:Pagina de autentificare Amazon

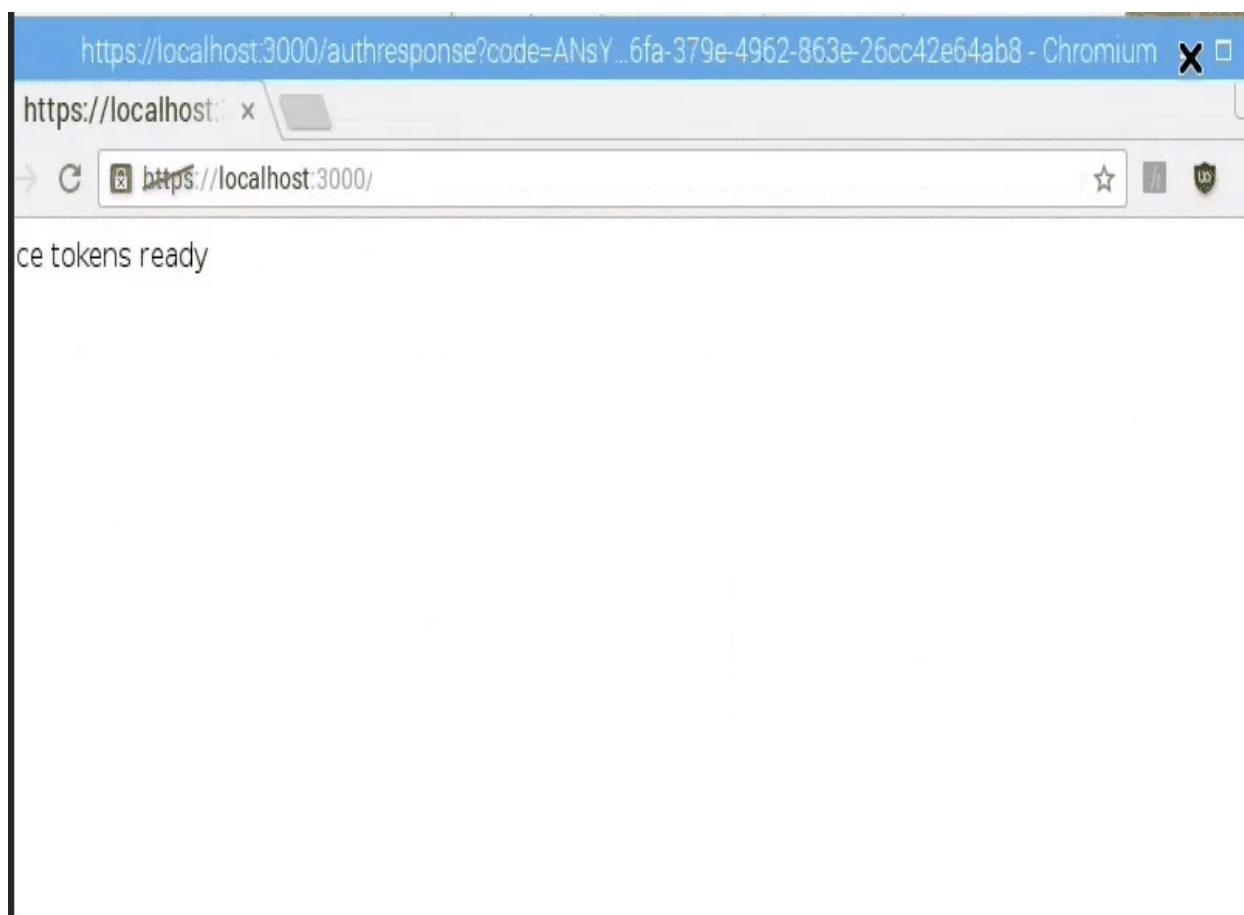


Figura 19: Confirmarea autentificării

Pasul 5 : Poniți Wake Word Engine

În cele din urmă, deschideți a treia fereastră Terminal (File > New Window). Aici, veți începe Wake Word Engine. Acest lucru face ca tu să poți spune "Alexa" pentru ca Raspberry Pi să înceapă să asculte. Aveți două opțiuni pentru software-ul word Wake, Sensory și KITT.AI. Ambele sunt gratuite, dar Sensory expiră după 90 de zile, deci să folosim KITT în schimb:

- Scriem în terminalul deschis `cd ~/Desktop/alexa-avs-sample-app/samples` și apoi apăsăm enter.
- Scriem `cd wakeWordAgent/src && ./wakeWordAgent -e kitt_ai`.

Asta e, DIY Echo care rulează acum. Mergeți și încercați să-i spuneți "Alexa". Ar trebui să auziți un bip indicând că asculta. Când auziți acel beep, adresați-i o întrebare cum ar fi "What's the weather" ?

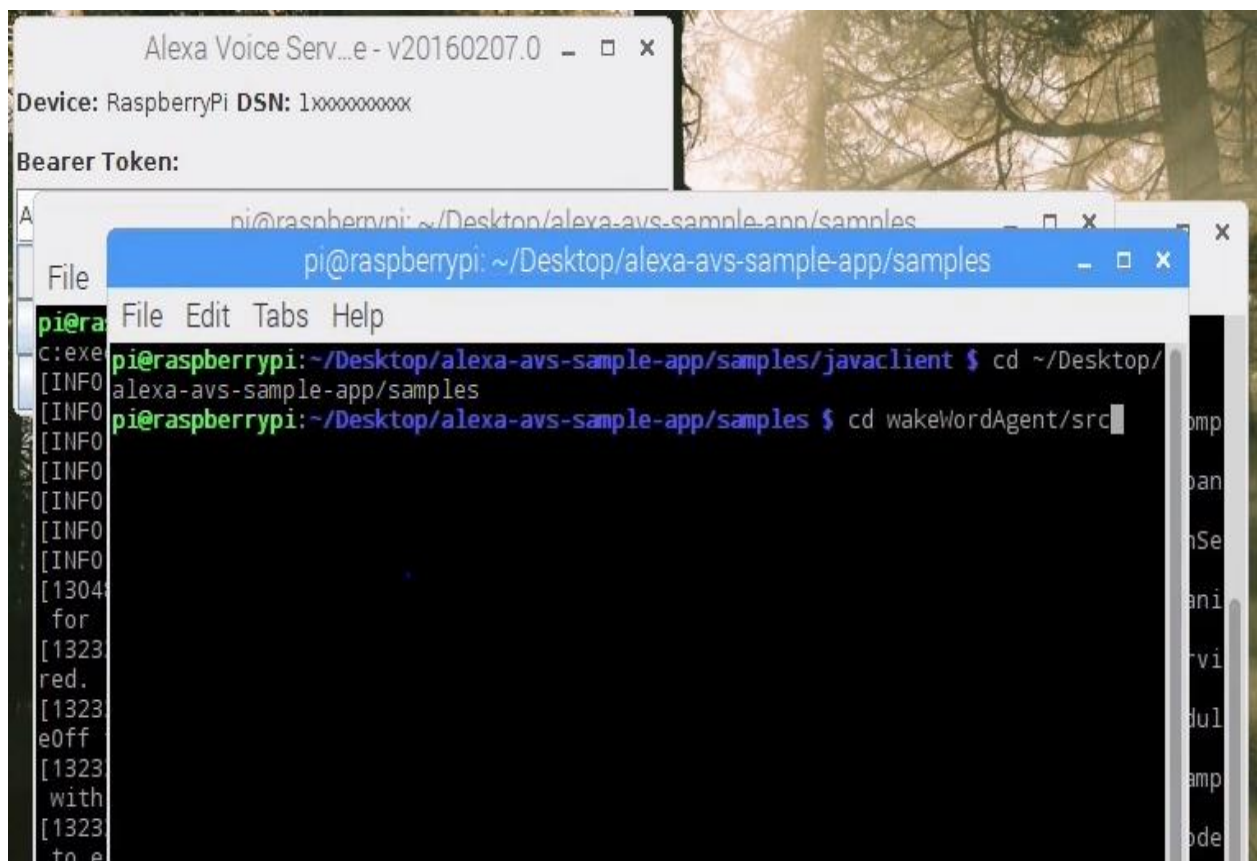


Figura 20:Lansarea scriptului care pornește recunoașterea

Pasul 6 : Îmbunătățirea microfonului și asigurarea faptului că Echo înțelege ce ai spus

În cele din urmă, în funcție de calitatea microfonului, este posibil să observați că are probleme în a vă auzi. În loc să țipați "Alexa" din toți plămânii, să mergem la linia de comandă pentru ultima oară.

- În linia de comandă tastăm alsamixer și apăsăm enter.
- Apăsăm F6 și selectăm un USB diferit. Vom folosi săgețile tastaturii pentru a selecta microfonul
- Din nou folosim săgețile pentru a crește volumul de captură
- Când ajungem la un nivel al volumului satisfăcător apăsăm exit și ieșim
- Tastăm sudo alsactl store și apăsăm enter pentru ca setările să devină permanente.

Acum, ar trebui să puteți să declanșați Echo-ul DIY, vorbind cu el ca un om normal, în loc să strigeți. De asemenea, puteți modifica volumul implicit aici dacă doriți. [1] [5]



Figura 21: Captură de ecran din setările microfonului

CAP.4 Instalarea Google Assistant pe Raspberry Pi.

I) Privire de ansamblu :

Acest proiect demonstrează cum să obțineți o recunoaștere naturală a limbajului și cum să îl conectați la Asistentul Google, utilizând kitul de voce AIY Projects. Împreună cu tot ceea ce Asistentul Google are deja, vă puteți adăuga propriile dvs. întrebări și perechi de răspunsuri. Toate într-un cub de carton mic, ușor de utilizat, alimentat de Raspberry Pi. De asemenea, puteți integra Asistentul Google în propriul hardware, urmând ghidurile oficiale SDK Google Assistant .

II)Lista de materiale.

Deschideți cutia și verificați dacă aveți toate componentele necesare în kit. De asemenea, veți avea nevoie de câteva instrumente pentru asamblare.

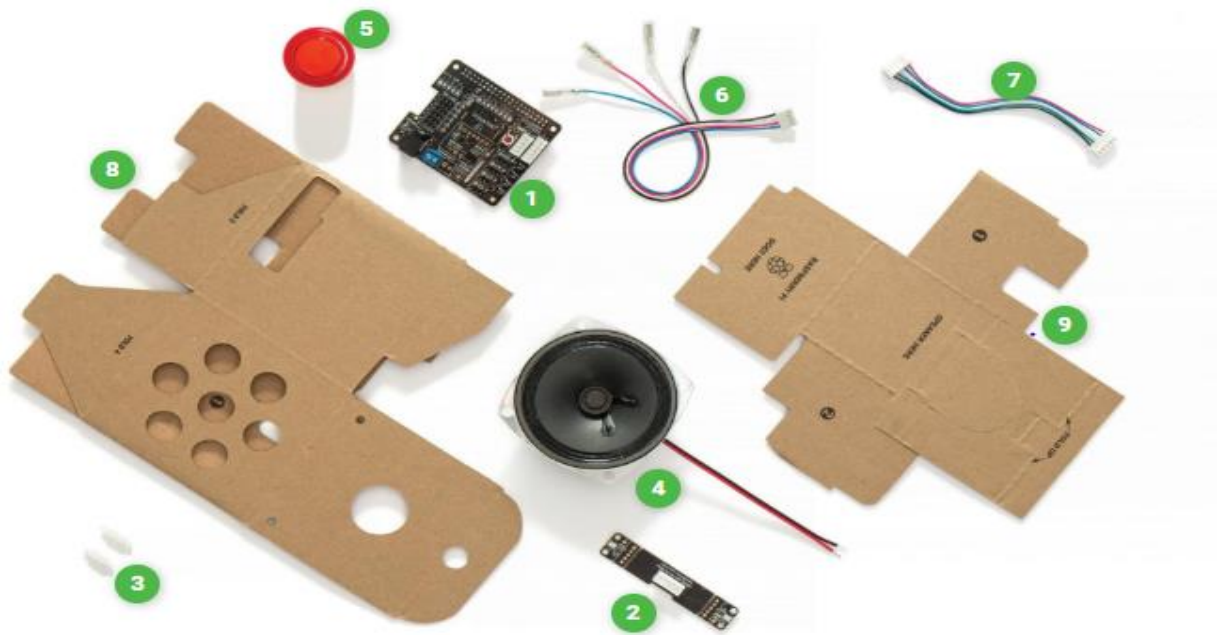


Figura 22 : Kit-ul Aiy project neasamblat

- 1.Placă de accesorii Voice HAT (× 1)
- 2.Placă de microfon Voice HAT (× 1)
- 3.Distanțiere de plastic (× 2)
- 4.3 "difuzor (cabluri atașate) (× 1)
- 5.Buton în stil arcade (× 1)
- 6.Cablu de buton cu 4 fire (× 1)
- 7.Cablu mamă cu 5 fire pentru placă (× 1)
- 8.Cutie de carton externă (× 1)
- 9.Cadru de carton intern (× 1)

NU ESTE INCLUS:

Raspberry Pi 3 (× 1)

Card SD (× 1)

Dimensiune "00" Șurubelniță Phillips (× 1)

Scotch bandă (× 1)

III) Ghidul de asamblare :

1) Scrieți imaginea pentru Voice Kit pe un card SD:

Va trebui să descărcați imaginea SD de voce folosind un alt PC. Ambii pași următori pot dura câteva minute până când PC-ul va termina treaba, așa că în timp ce așteptați, începeți să "Asamblați hardware-ul" în pasul următor.

a. Descărcați Voice Kit SD image

b. Scrieți imagine pe cardul SD utilizând o aplicație de scriere specializată (Etcher.io)

2. CONECTAREA ȘI PORNIREA DISPOZITIVULUI

2.1. CONECTEAZĂ PERIFERIILE



Figura 23: Asamblare tuturor componentelor proiectului Aiy

Acum, când cutia este asamblată, conectați perifericele la:

1 tastatură USB

2 mouse USB

3 monitor HDMI

2.2. Porniți dispozitivul:

Introduceți cardul SD (cel cu imaginea SD de pe dispozitivul de voce) în slotul de pe partea inferioară a plăcii Raspberry Pi. Slotul cardului SD trebuie să fie accesibil printr-o decupare furnizată în cutia de carton. Cu cardul SD la locul lui și perifericele conectate, conectați alimentarea electrică, iar Raspberry Pi va începe să se încarce. Dacă nu vedeți nimic pe monitor sau vedeți "Eroare de sintaxă Openbox", verificați ghidul de depanare din anexă.



Figura 24: Inserarea cardului de memorie în Raspberry Pi

2.3. CONECTAREA LA INTERNET:

Faceți clic pe pictograma de rețea din colțul din dreapta sus al desktop-ului Raspberry Pi. Alegeți punctul de acces WiFi preferat.

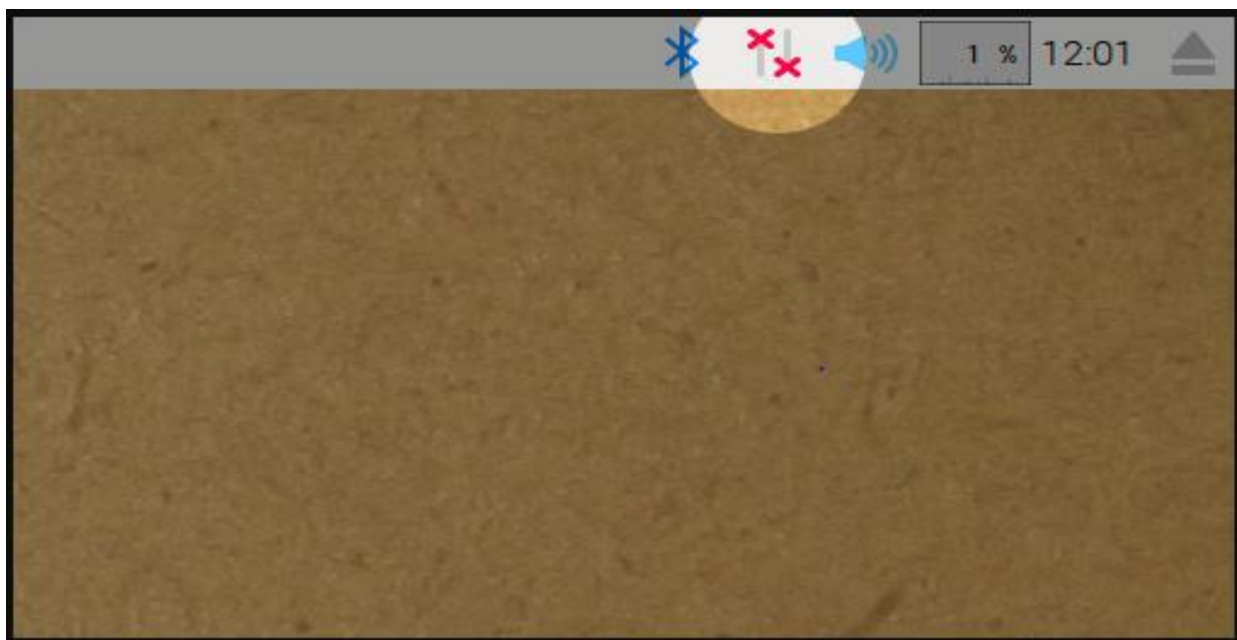


Figura 25: Conectarea la internet a Raspberry Pi-ului

3. Verificarea funcționalității:

După ce a fost lansat, LED-ul roșu de pe Raspberry Pi lângă conectorul de alimentare trebuie să fie aprins. Dacă nu, verificați ghidul de depanare.

3.1. VERIFICAREA AUDIO

1) Acest script verifică dacă funcționează corect componentele de intrare și ieșire audio de pe placa de accesorii HAT. Faceți dublu clic pe pictograma Check Audio de pe desktop. Când faceți clic pe script, acesta va rula prin fiecare pas enumerat mai jos. Notă: unii dintre pași necesită introducere vocală, ceea ce vi se va solicita, așa că urmăriți îndeaproape!



Figura 26: Lansarea scriptului de verificare a voice Hat-ului(microfon și boxă)

2) Urmăți împreună cu scriptul și dacă totul funcționează corect, veți vedea un mesaj care spune că sunetul pare să funcționeze

Dacă vedeți un mesaj de eroare, urmați detaliile mesajului pentru a rezolva problema și încercați din nou.

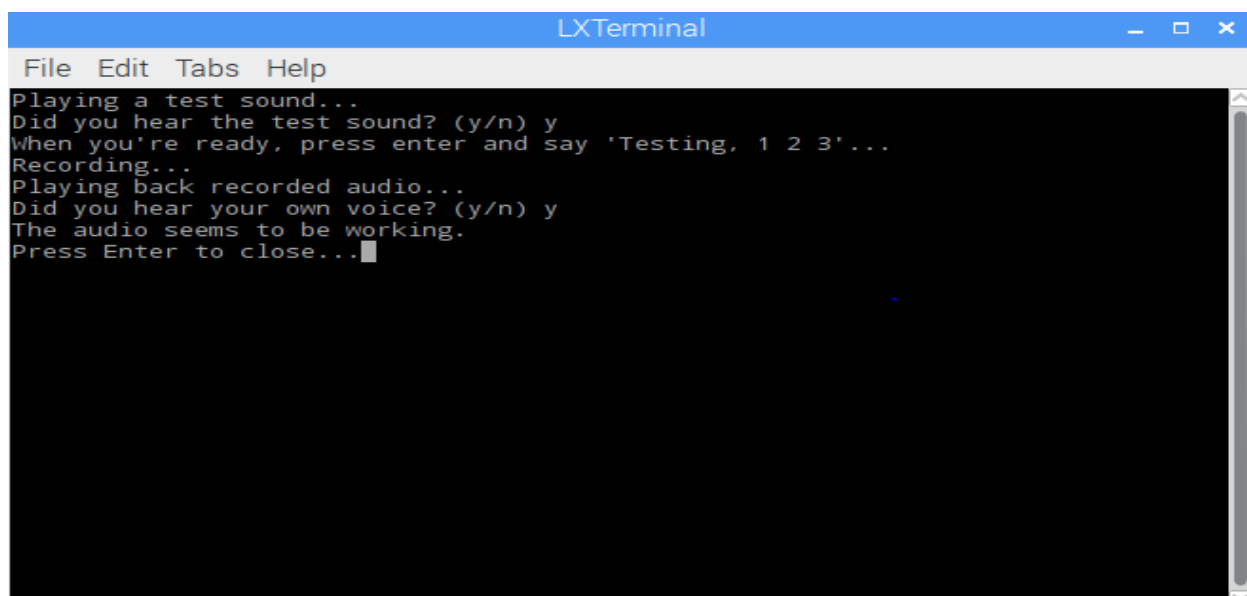


Figura 27: Rezultatul favorabil în urma verificării

3.2. VERIFICARE WIFI

3)

Acest script verifică dacă WiFi-ul dvs. este configurat și funcționează corect pe placa Raspberry Pi. Faceți dublu clic pe pictograma Check Wi-Fi de pe desktop. Când faceți dublu clic pe script, va verifica dacă Raspberry Pi este conectat la internet prin WiFi.

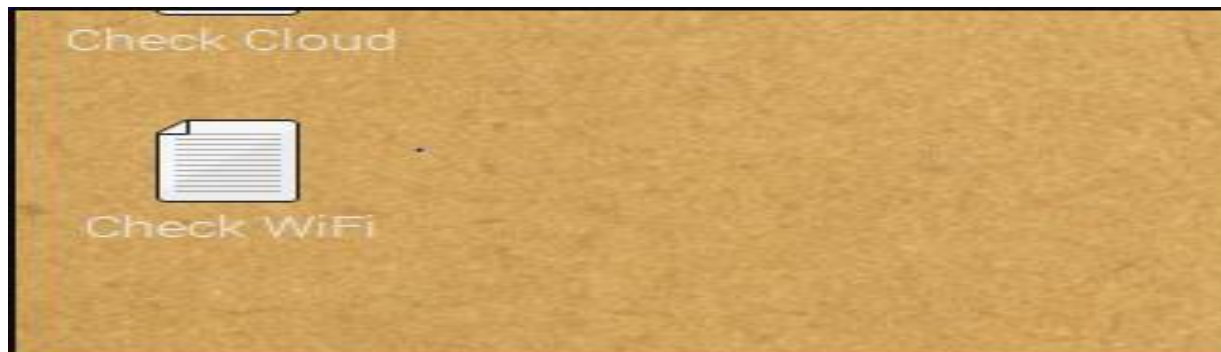


Figura 28: Lansarea scriptului de verificare a conexiunii la internet

4)

Dacă totul funcționează corect, veți vedea un mesaj care spune că conexiunea WiFi pare să funcționeze. Dacă vedeți o eroare, dați clic pe pictograma de rețea din dreapta sus și verificați că sunteți conectat (ă) la un punct de acces valabil.

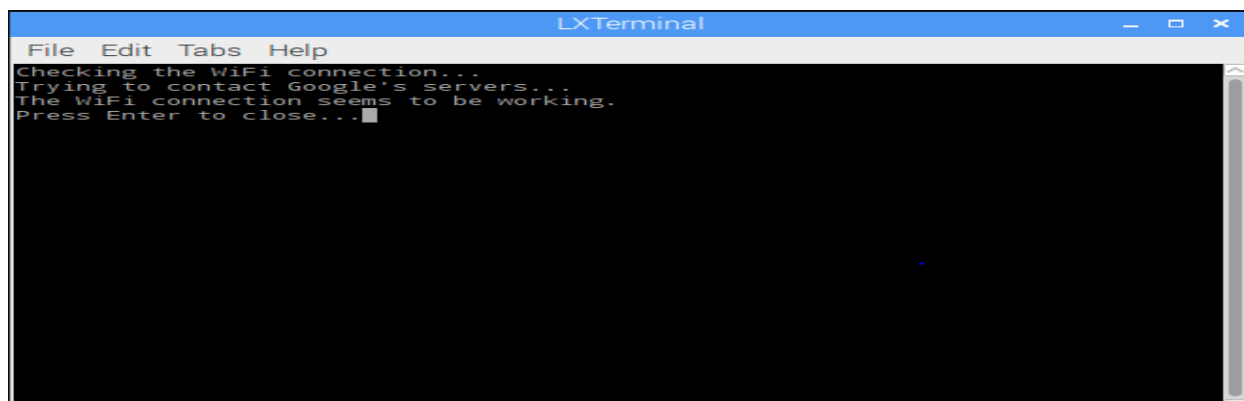


Figura 29 : Rezultatul favorabil în urma rulării scriptului

4)ANEXA:

Ghidul de depanare:

- A. Un LED roșu pe Raspberry Pi lângă conectorul de alimentare trebuie să se aprindă. Dacă nu, deconectați alimentarea, deconectați conectorul de la microfon și reporniți alimentarea. Dacă se aprinde după pornire fără microfon, atunci placa de microfon poate fi defectă.
- B. Lampa din buton nu se aprinde până când nu executați o demonstrație, așa că nu vă faceți griji că este oprit
- C. Dacă nu vedeți nimic pe monitor, asigurați-vă că cablurile HDMI și cablurile de alimentare sunt complet introduse în Raspberry Pi.

- D. Dacă vedeți "Eroare de sintaxă Openbox", va trebui să rescrieți imaginea pe cardul SD și să încercați din nou să porniți dispozitivul.

IV) Ghidul Utilizatorului

1. SETAREA DISPOZITIVULUI DUMNEAVOASTRĂ.

1.1. CONECTAREA PLATFORMEI GOOGLE CLOUD

Pentru a încerca API-ul Google Assistant, trebuie să vă conectați mai întâi la Google Cloud Platform (GCP) și apoi să activați API-ul.

Conectați-vă la GCP :

1) Folosind dispozitivul dvs. de recunoaștere a vocii, deschideți un browser de internet și mergeți la Consola cloud.

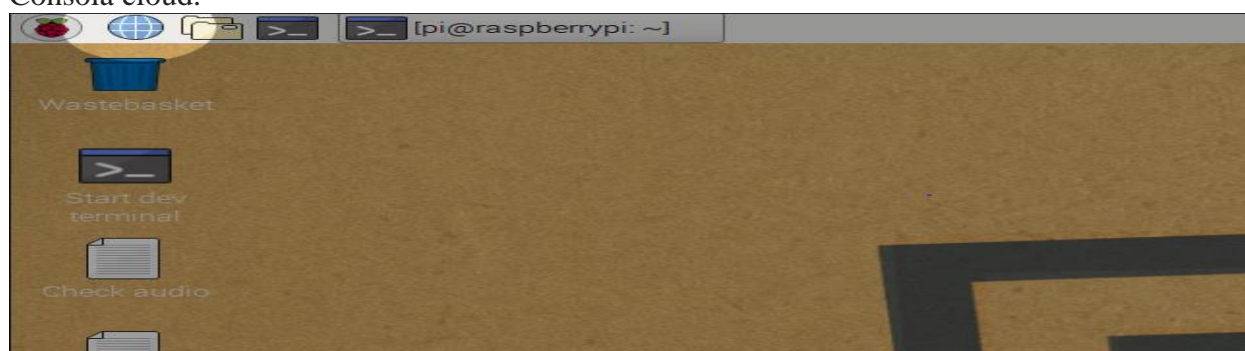


Figura 30: Deschiderea browser-ului de internet

NU AȚI MAI UTILIZAT PLATFORMA GOOGLE CLOUD NICIODATA :

Utilizați contul dvs. Google pentru a vă conecta. Dacă nu aveți unul, va trebui să creați unul.

Încercarea API-ului Google Assistant este gratuită pentru uz personal.

Creați un proiect: GCP utilizează proiectele pentru a organiza lucrurile. Creați unul pentru caseta de recunoaștere vocală.

2) În Consola cloud, faceți clic pe butonul vertical din dreapta paginii "Google Cloud Platform". Din meniul derulant, faceți clic pe Create Project.

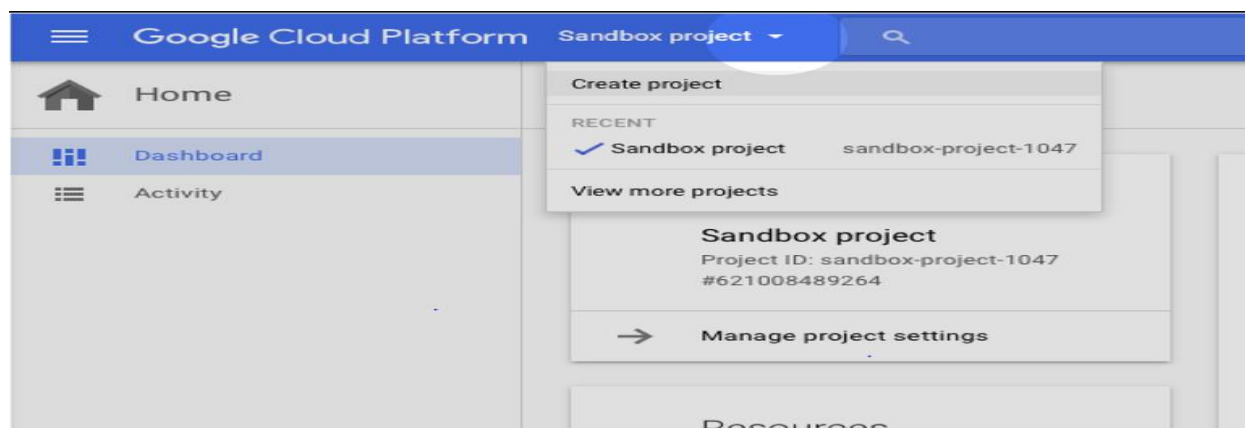


Figura 31: Deschiderea interfeței Google Cloud Platform și selectarea unui proiect nou

3) Introduceți un nume pentru proiect.

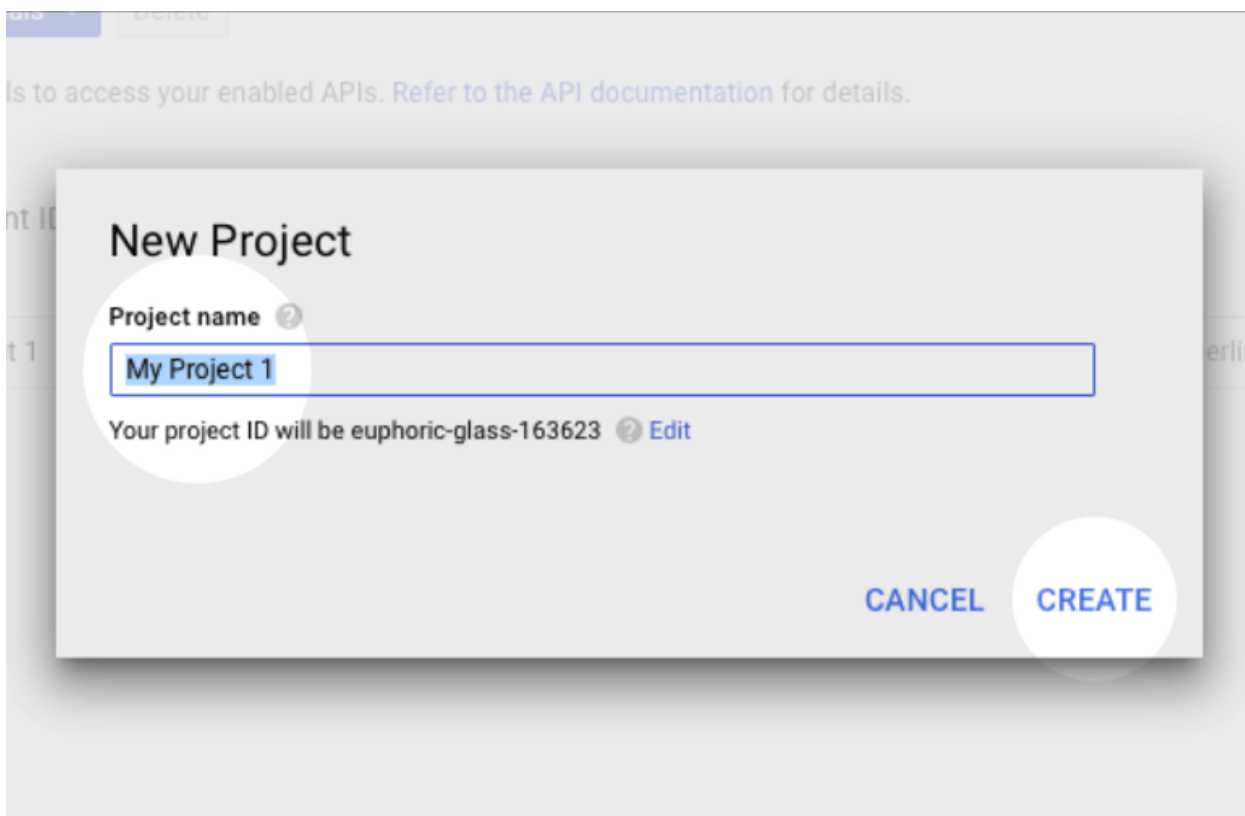


Figura 32: Crearea proiectului în Google Cloud Platform cu numele dorit

- 4) După crearea proiectului, asigurați-vă că meniul derulant are noul nume de proiect afișat

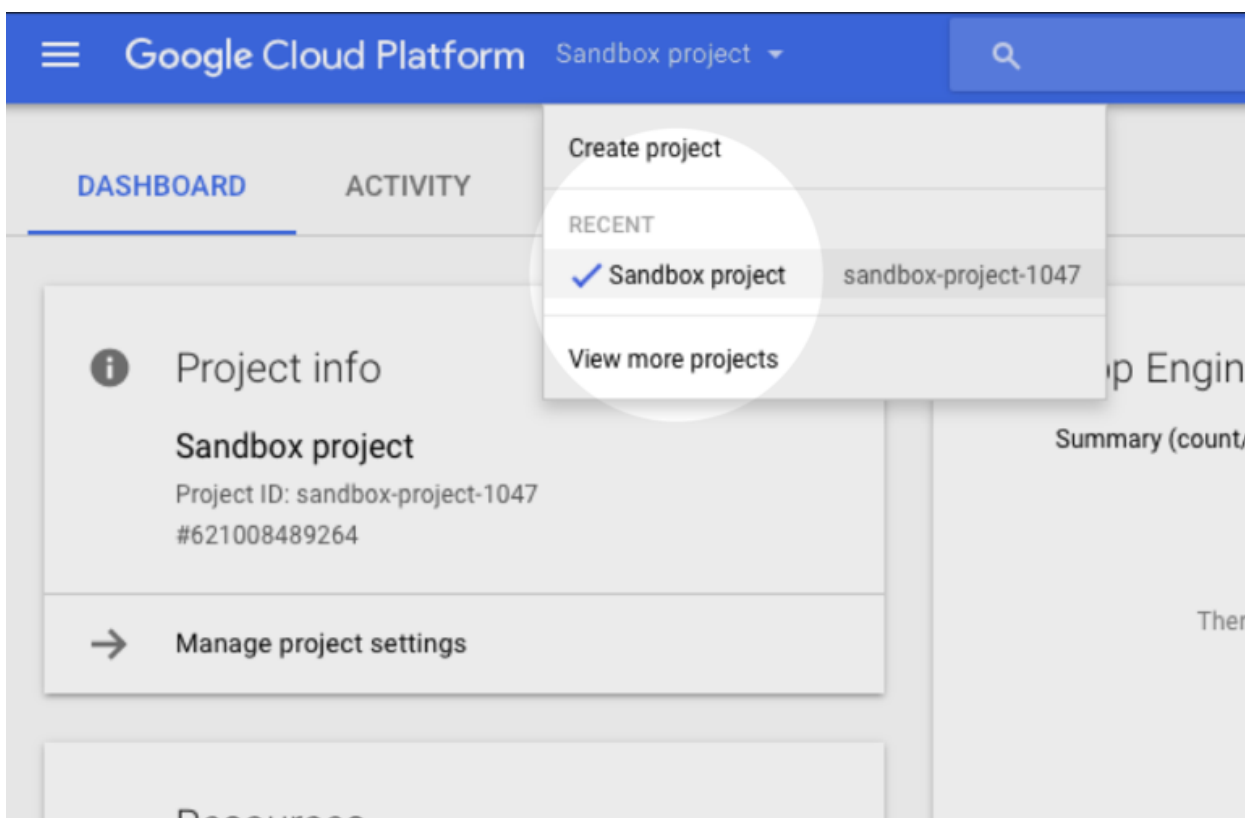


Figura 33: Verificarea numelui proiectului

- 5) În Consola cloud, activați "API-ul Google Assistant".

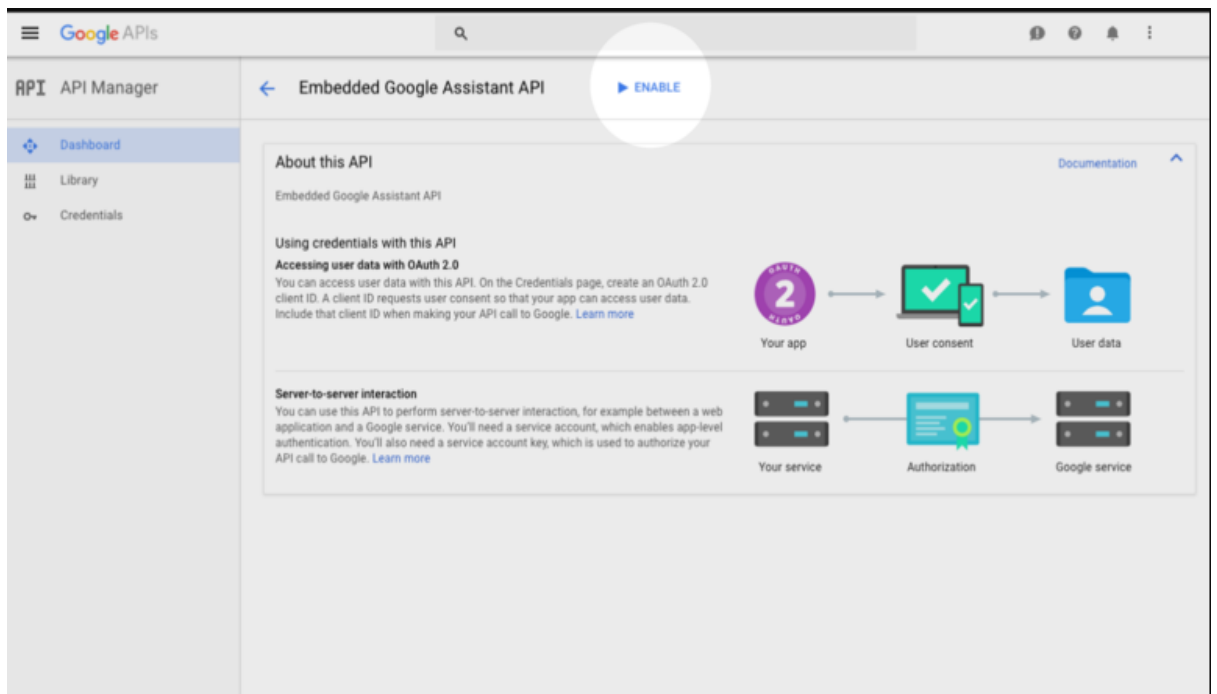


Figura 34: Activarea API-ului Google Assistant

- 6) În Consola cloud, creați un client OAuth 2.0 accesând APIs & Services> Credentials

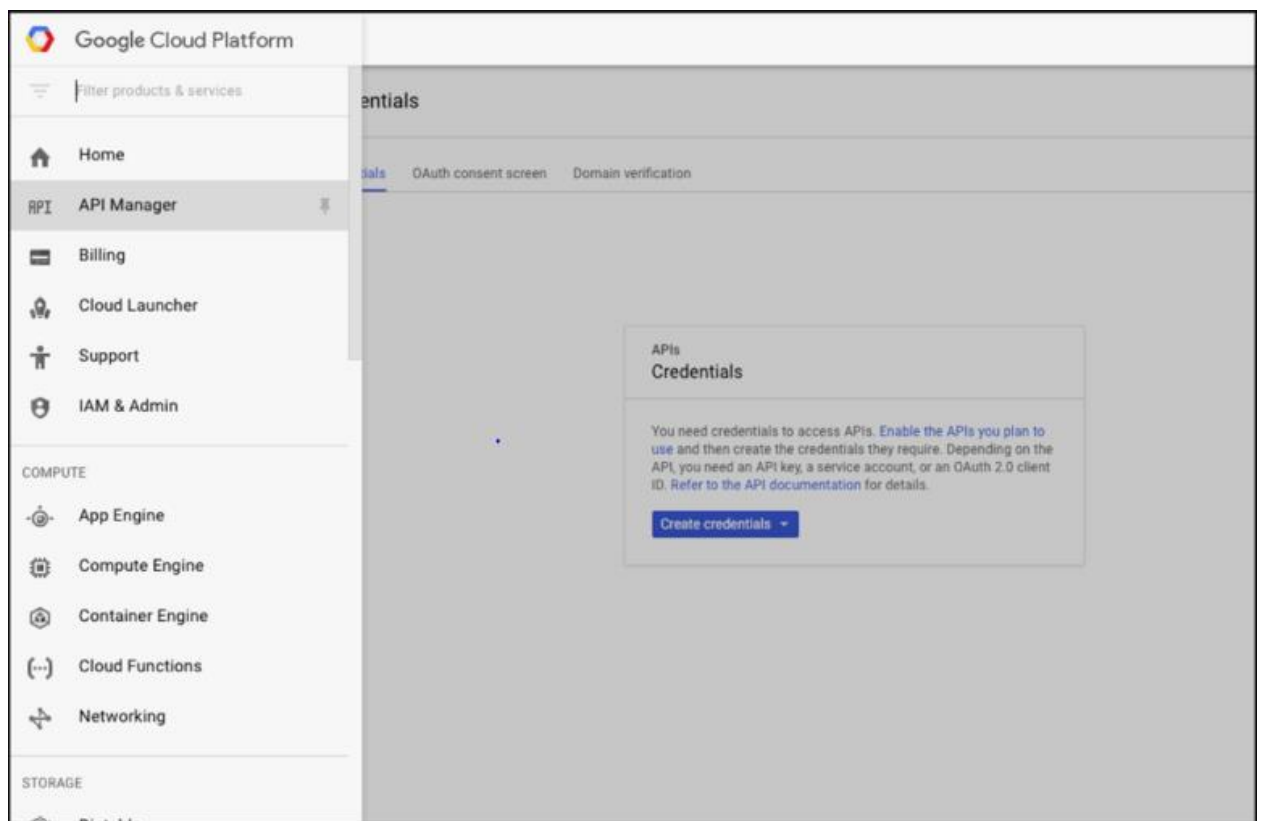


Figura 35: Crearea clienților pentru autentificare

- 7) Faceți clic pe Creați acreditări și selectați ID client OAuth

Dacă aceasta este prima dată când creați un ID de client, va trebui să configurați ecranul de consimțământ făcând clic pe Configurați ecranul de consimțământ. Va trebui să denumiți aplicația (acest nume va apărea în etapa de autorizare).

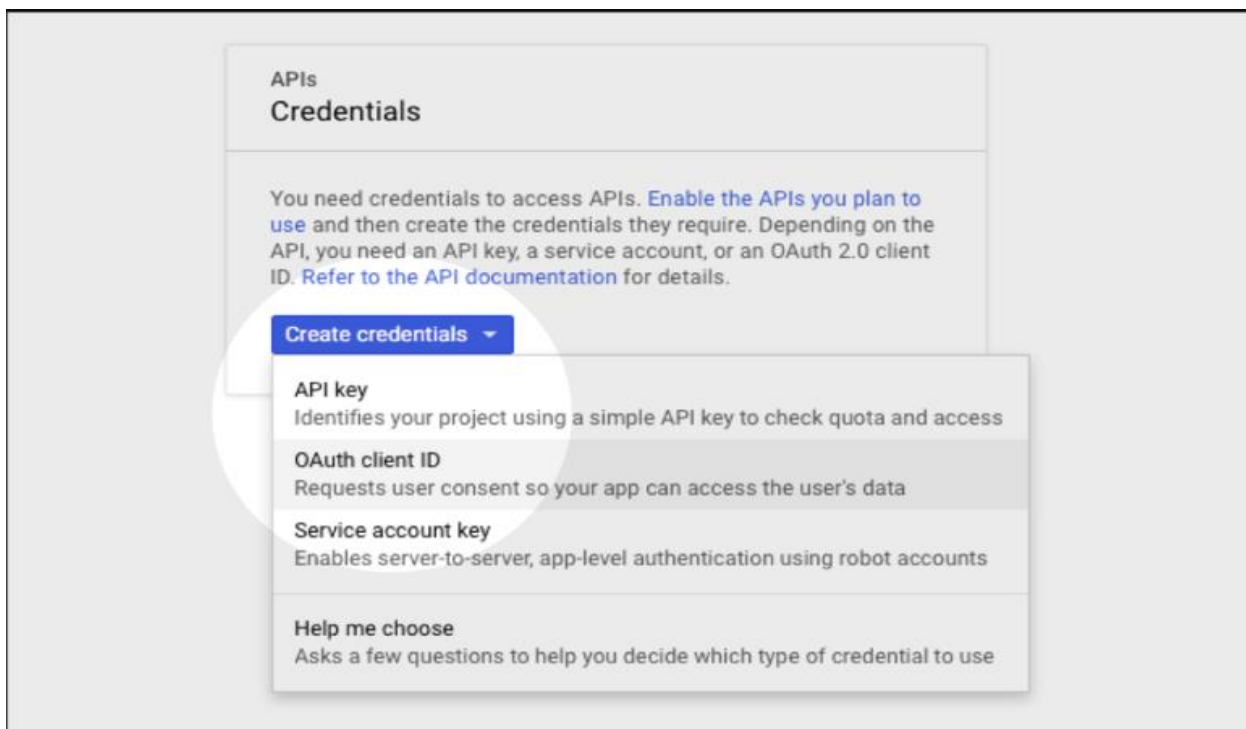


Figura 36: Selectarea OAuth id-ului pentru credentials

8) Selectați Altele, introduceți un nume pentru a vă ajuta să vă amintiți acreditările, apoi faceți clic pe Creați.

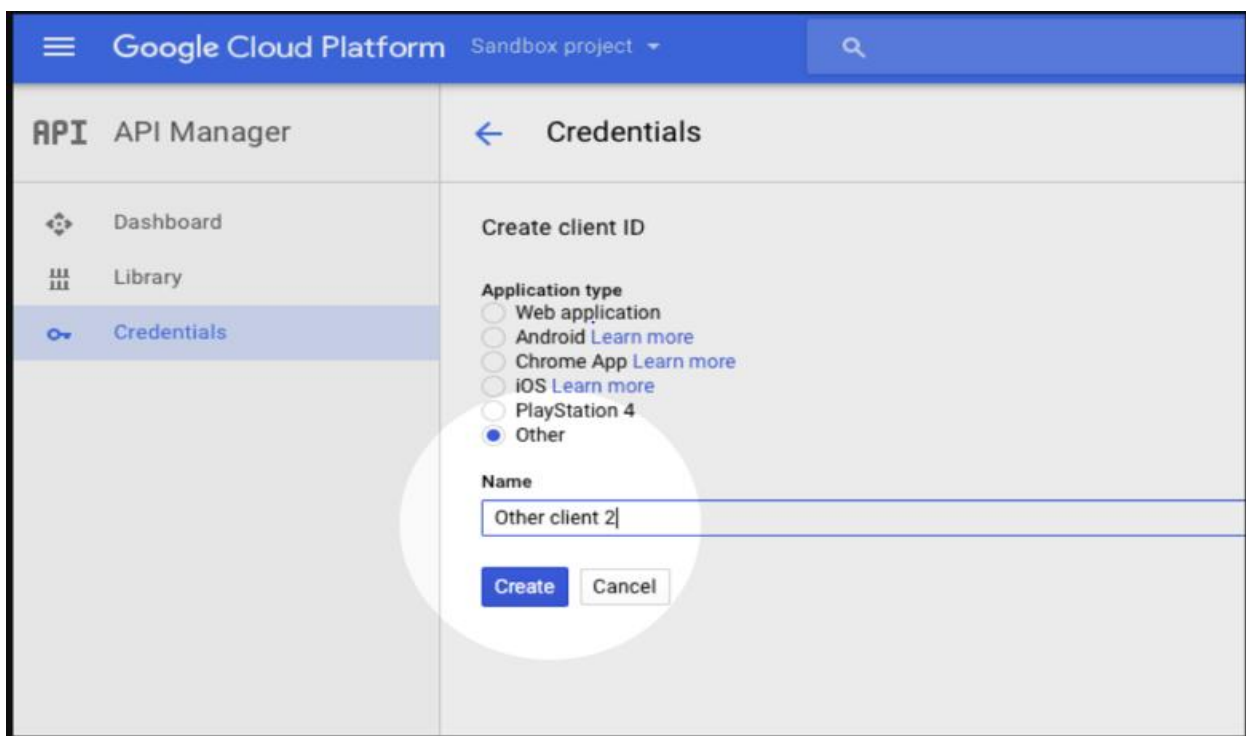


Figura 37: Selectăm Other și selectăm numele pentru credentials

9) Va apărea o fereastră de dialog. Faceți clic pe OK. În lista de acreditare, găsiți-vă noile acreditări și faceți clic pe pictograma de descărcare (pictograma Descărcare) din partea dreaptă.

Notă: dacă nu vedeți pictograma de descărcare, încercați să extindeți lățimea ferestrei browserului sau să micșorați.

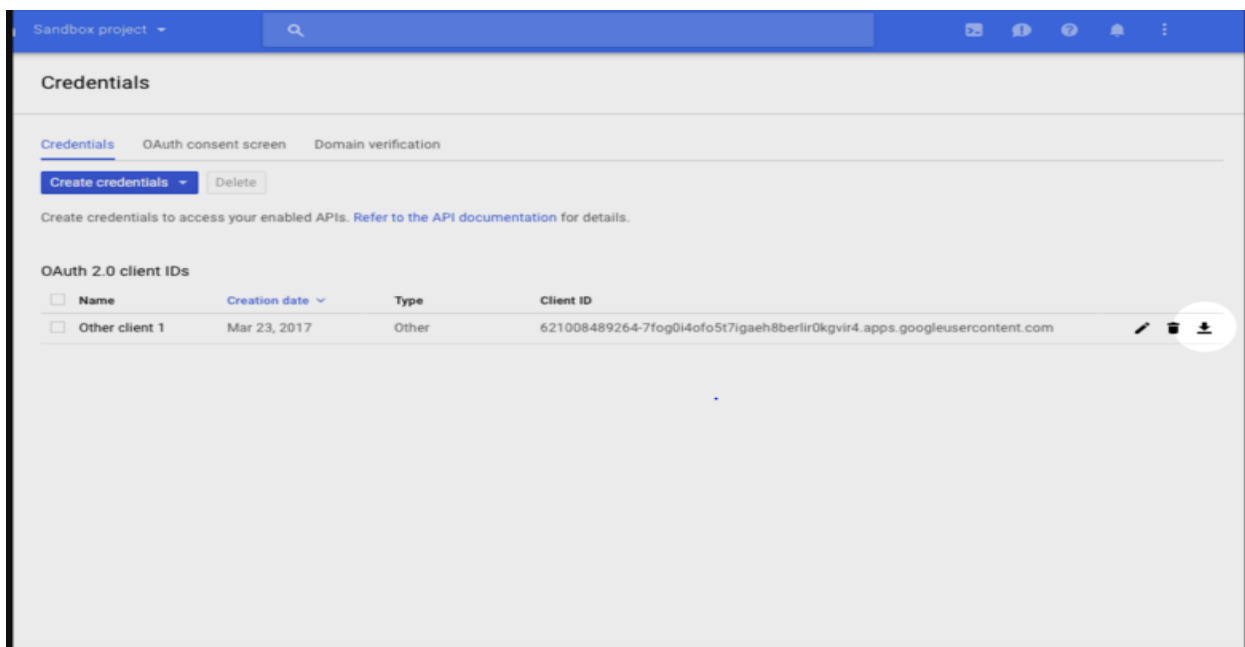


Figura 38:Fereastră de dialog din care putem descărca fișierul de credentials cu extensia json

10) Găsiți fișierul JSON pe care tocmai l-ați descărcat (client_secrets_XXXX.json) și redenumi-l la asistent.json. Apoi îl mutați la /home/pi/assistant.json

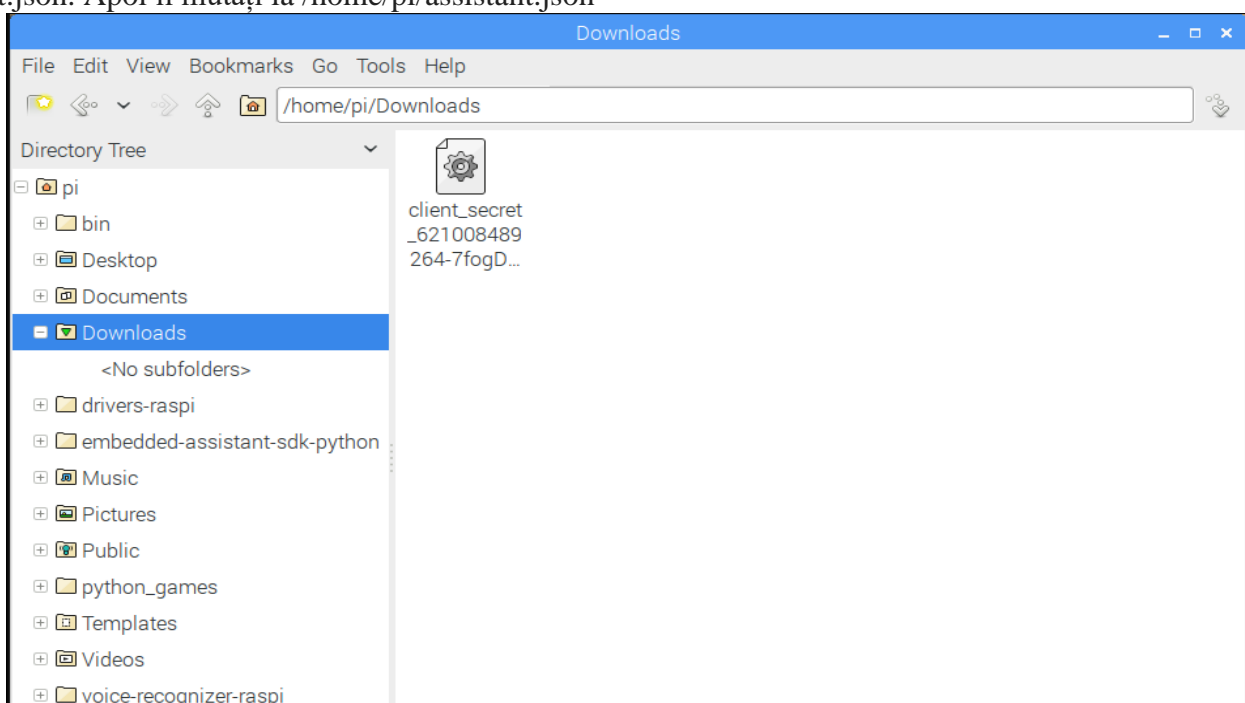


Figura 39: Fișierul json descărcat ce urmează să fie redenumit și mutat

11) Accesați panoul de control al activității. Asigurați-vă că vă conectați cu același cont Google ca înainte. Activați următoarele: Activitatea web și a aplicațiilor, Informație despre dispozitiv Activitatea voce și audio

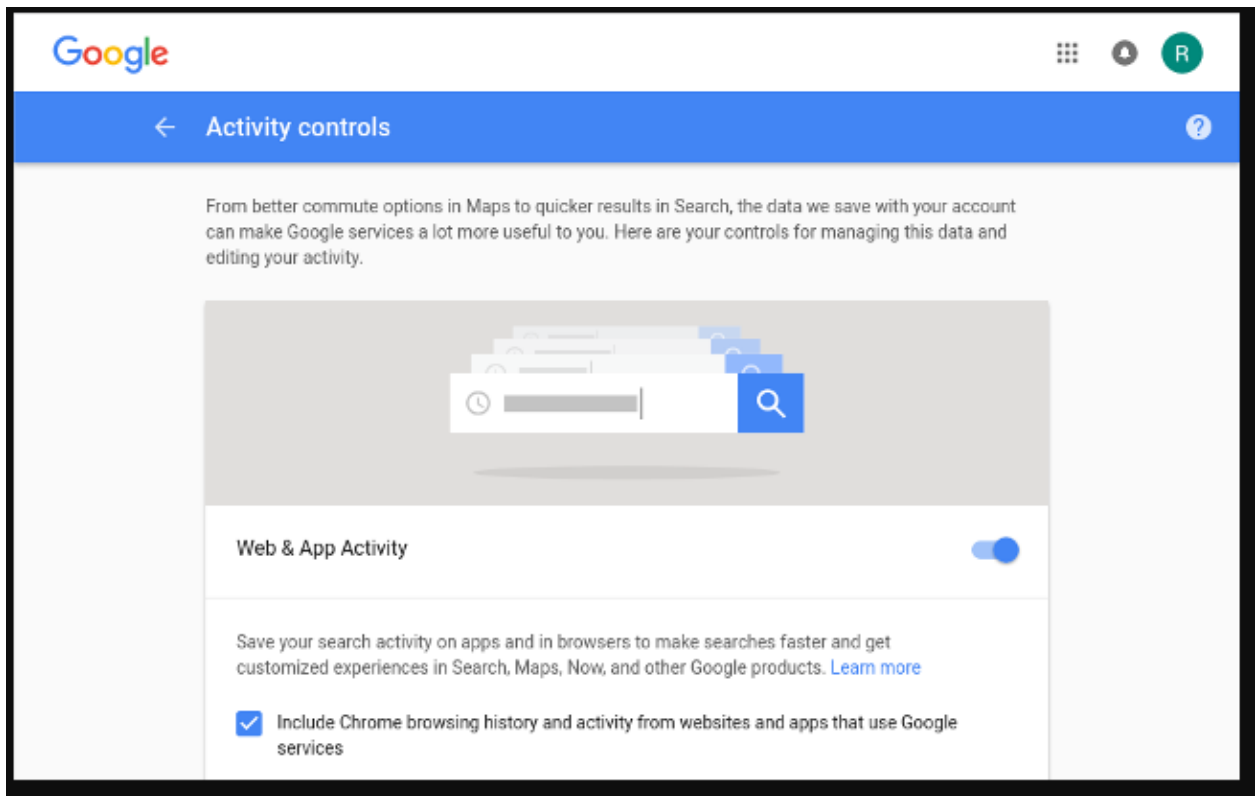


Figura 40: Activarea Web&App Activity

12) Sunteți gata să îl porniți: urmați instrucțiunile de pornire manuală din secțiunea Utilizarea dispozitivului de mai jos .

13) Puteți, de asemenea, SSH de pe alt computer. Va trebui să utilizați ssh -X pentru a gestiona autentificarea prin browser atunci când porniți exemplul pentru prima dată.

Autorizați accesul la API-ul Google Assistant, când vi se solicită acest lucru

14) Asigurați-vă că urmați instrucțiunile de pornire manuală prima dată - dacă executați ca serviciu, nu vi se va solicita autorizare. Încercați un exemplu de interogare precum "ce este în calendarul meu?" - și asistentul ar trebui să răspundă!

2. Folosirea device-ului:

Sunt oferite trei aplicații demo care prezintă recunoașterea vocală și Asistentul Google cu diferite capabilități. Acestea pot fi folosite ca șabloane pentru a crea propriile aplicații.

Când rulează o aplicație demo, LED-ul din interiorul butonului arcade și LED-ul din centrul Hat Voice vor emite pulsuri la fiecare câteva secunde. Dacă nu vedeți pulsul LED, verificați ghidul de depanare.

DEMO APP	DESCRIPTION	RASPBERRY PI SUPPORTED
assistant_library_demo.py	Showcases the Google Assistant Library and hotword detection ("Okay, Google").	2b, 3b
assistant_grpc_demo.py	Showcases the Google gRPC APIs and button trigger.	2b, 3b, Zero W
cloudspeech_demo.py	Showcases the Google Cloud Speech APIs, button trigger, and custom voice commands.	2b, 3b, Zero W

Figura 41: Fișierele python disponibile pentru utilizare

2.1 PORNEȘTE ASSISTANT LIBRARY DEMO APP

Pentru ca dispozitivul să înceapă să acționeze în calitate de asistent Google, la fel ca Google Home, porniți aplicația demo de asistare a bibliotecii făcând dublu clic pe "Start terminal dev" pe Desktop și introducând:

```
src / exemple / voce / assistant_library_demo.py
```

Aplicația bibliotecă asistentă are încorporat detecția cuvintelor cheie. Pentru a începe o conversație cu Asistentul Google, spuneți "Ok, Google" sau "Hei Google". Când ați terminat, apăsați Ctrl-C pentru a termina aplicația.

2.2 PORNEȘTE ASSISTANT GRPC DEMO APP

Faceți dublu clic pe "Start dev terminal" de pe desktop și introduceți:

```
src / exemple / voce / assistant_grpc_demo.py
```

Spre deosebire de demo-ul asistent al bibliotecii, această demonstrație nu acceptă detectarea cuvintelor cheie.

Pentru a solicita o întrebare din partea Asistentului Google, apăsați butonul arcade și vorbiți. Când ați terminat, fie apăsați butonul arcade și spuneți "la revedere", fie pur și simplu apăsați pe Ctrl-C pentru a termina aplicația.

2.3 PORNEȘTE CLOUD SPEECH DEMO APP

Demo-ul vorbesc prin cloud utilizează API-urile Google Cloud Speech. Dacă nu aveți nevoie de conversațiile oferite de Asistentul Google, acest lucru este util pentru crearea unei aplicații proprii pentru recunoașterea comenzilor vocale. [4]

src / exemple / voce / cloudspeech_demo.py

CAP.5 Google AIY Voice Kit

Termenul "AIY" este un "tip" de acronim pentru " Artificial Intelligent Do It Yourself", o inițiativă a companiei Google de a oferi medii de dezvoltare cu costuri reduse pentru a experimenta tehnologii de inteligență artificială de vârf.

Cu Voice Kit Google a făcut posibil ca oamenii pasionați să adauge cu ușurință recunoașterea vocii și feedback-ul vocal la proiectele lor.

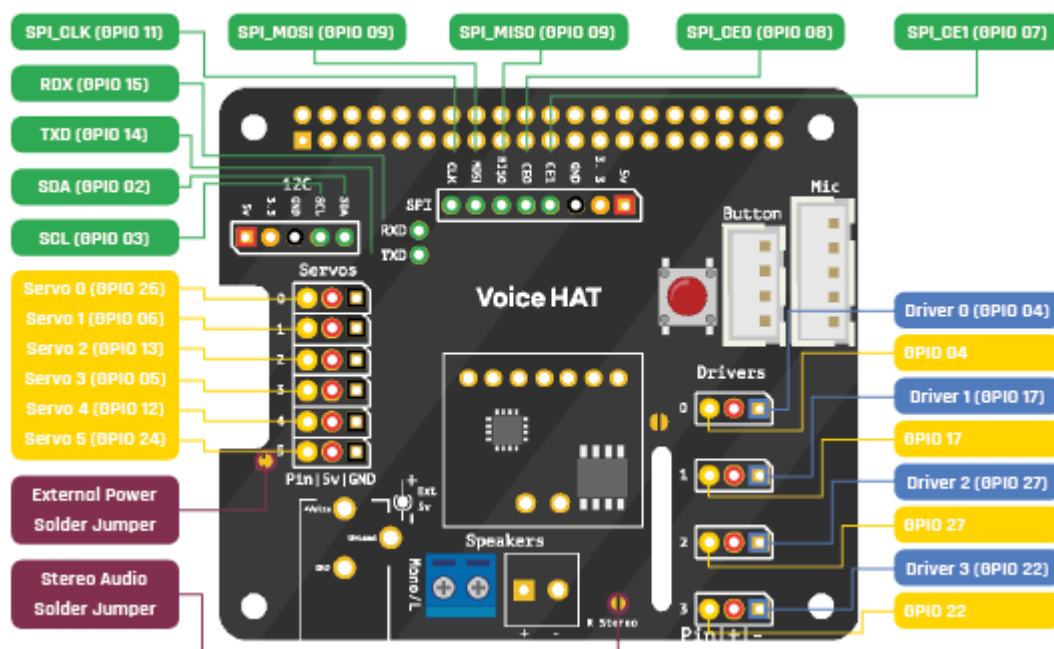


Figura 42: Reprezentarea pinilor Voice Hat-ului

Componenta hardware principală din setul AIY Voice Google este AIY Voice HAT. Ca orice alt HAT, aceasta este o placă care se atașează la Raspberry Pi prin GPIO cu 40 de pini.

HAT este locul în care veți atașa dispozitive externe pe care le puteți controla apoi cu vocea dumneavoastră. Ca atare, poate fi mai ușor să lucrați dacă eliminați întreaga combinație de Raspberry Pi și HAT din limita cutiei de carton. Aceasta va face cablarea mult mai ușoară.

Deoarece HAT este componenta hardware principală din kitul AIY Voice Kit, ar trebui să ne uităm mai atent la acesta. În multe privințe, AIY Voice HAT este într-adevăr o placă de sunet foarte deșteaptă. Este nevoie de o intrare de pe placa de microfon pentru a trimite informația până la Google Cloud. Nu recunoaște sau interpretează vocea de la sine. Când răspunsul este returnat de Google, AIY Voice HAT ia răspunsul respectiv și îl convertește în sunet. Amplificatorul audio încorporat este utilizat pentru a conduce difuzorul. AIY Voice HAT este de fapt capabil de ieșire

stereo, dar în mod implicit acționează doar difuzorul stâng în modul monofonic. AIY Voice HAT are, de asemenea, conexiuni la Raspberry Pi GPIO expuse pe un număr de pini pentru atașarea dispozitivelor externe.

În plus față de conectorul cu 40 de pini Raspberry Pi GPIO, Google AIY Voice HAT are următoarele conexiuni:

1. Microfon - Un conector cu 5 pini pentru a atașa placa microfonului stereo.
2. Buton de apăsare - Un conector cu 4 pini pentru butonul cu LED integrat.
3. Ieșiri (4) - Fiecare dintre acestea este o conexiune cu 3 pini.
4. Mono / Stereo Jumper - Vă permite utilizarea canalului audio din dreapta pentru ieșirea stereo.
5. Al doilea difuzor - o conexiune pentru difuzorul din dreapta când funcționează în modul stereo.
6. Alimentare externă - Conexiuni pentru o sursă de alimentare externă, utilizată atunci când folosiți dispozitive cu curent înalt care ar fi prea mari pe sursa de alimentare cu 5 volți de la Raspberry Pi.
7. Alimentator extern - Protecția trebuie tăiat (rupt) pentru a utiliza sursa externă de alimentare.
8. Servo Motoare (6) - Fiecare dintre aceste conexiuni cu 3 pini poate fi utilizată pentru a conduce un servomotor. Ele pot fi, de asemenea, folosite pentru a acționa LED-urile și alte dispozitive cu curent redus.
9. Receptorul I2C - o conexiune la magistrala Raspberry Pi I2C pentru atașarea dispozitivelor externe I2C.
10. Serial RX & TX - O conexiune serială la Raspberry Pi
11. SPI Bus - O conexiune la magistrala SPI pentru conectarea perifericelor SPI.

AIY Voice HAT nu vine cu fire lipite pentru aceste conexiuni, ci doar găuri goale pe placa de circuite imprimate.

Servomotoare

Începând din partea stângă, există 6 seturi cu 3 pini care sunt etichetate ca "Servos". Servomotorul dorit a fost posibil prin utilizarea modulului PWM de la Raspberry Pi 3. Fiecare set are un pin GPIO, conexiune 5V și GND. Pinul GPIO nu se conectează direct la antetul Raspberry Pi 3, ci mai degrabă prin rezistoare de limitare a curentului de 220Ohm (etichetate R1-R6).

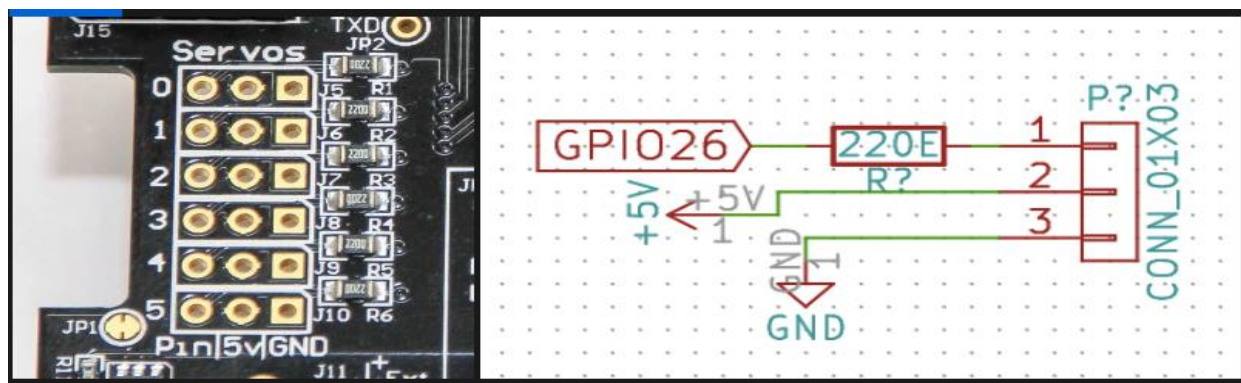


Figura 43: Pinii dedicați pentru servomotoare de pe Voice Hat

Sursa de Alimentare

Mai jos de acestea sunt dispozitive etichetate Q5 și Q6 care fac parte dintr-un circuit de selecție a sursei de alimentare. Unde Q5 se aprinde numai când tensiunea de intrare este mai mare decât 5 V de la portul USB. Un simplu comparator ar trebui să facă acest lucru (folosind LM393 pentru referință).

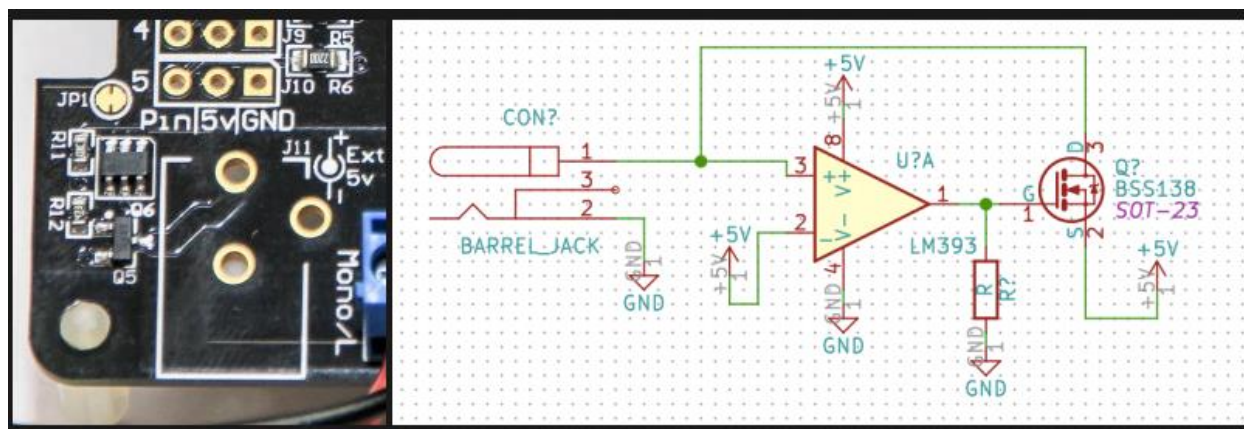


Figura 44: Pinii dedicați pentru sursă de alimentare externă de pe Voice Hat

Interfețe de comunicare

Mai sus de "Servo" este J15 etichetat I2C care se conectează direct la pinii Raspberry Pi 3. Asta înseamnă că acestea nu ar trebui să fie conectate la nimic cu 5V sau rezistențe de pull-up. Acestea nu sunt utilizate pe placă. Chiar lângă acesta se află partea de SPI și cei 2 pini UART.

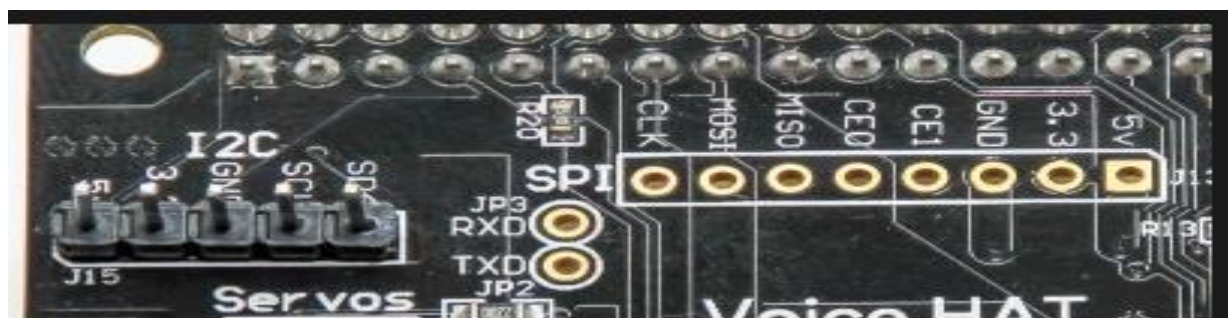


Figura 45: Pinii dedicați interfețelor de comunicație de pe Voice Hat

DAC și EEPROM

Puțin mai jos ajungem la circuitul integrat cu 16 pini marcat cu "AKK BDQ". Acesta este Maxim MAX98357A, care este un DAC I2S cu un amplificator de clasa D. Acesta este conectat direct la difuzorul, deoarece există doar o ieșire, acesta poate fi mono sau stereo.

Lucrul interesant este prezența JP6, care pare să aibă toate conexiunile I2S de la Maxim MAX98357A și alte câteva linii selectate. Combinat cu cele două intrări care se conectează la cea de-a doua ieșire a difuzoarelor, ar putea fi posibil să fie destinat pentru o altă placă de Maxim MAX35357A pentru a obține sunet stereo.

Alături de DAC este un SSOP cu 8 pini, care este un eEPROM I2C. Nu este conectat la pinii I2C despre care am vorbit mai devreme, ci mai degrabă la pini 27 și 28 din pinout-ul Raspberry Pi 3. EEPROM-ul deține informațiile producătorului de bord, configurarea GPIO și un lucru numit "device tree" - în esență, o descriere a hardware-ului atașat, care permite Linuxului să încarce automat driverele necesare.

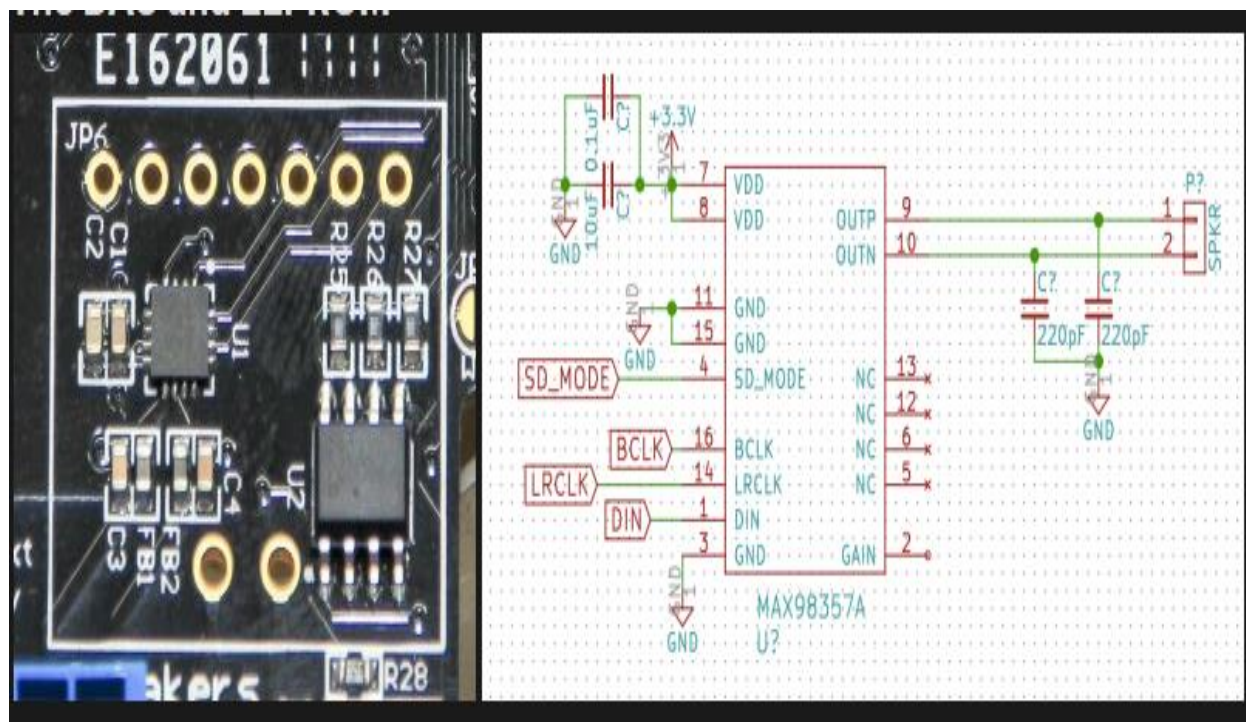


Figura 46: Memoria EEPROM și convertorul DAC

Drivere

Mergând spre dreapta, găsim 4 titluri marcate "Drivere". Acestea sunt circuite MOSFET pentru controlul unor relele.

MOSFET-urile pot suporta curenți de până la 500 mA fiecare, dar GPIO-urile sunt disponibile și pentru utilizare directă. Sarcinile care trebuie conduse trebuie să fie conectate între pinii marcați cu "+" și "-". Pin-ul din stânga este un acces direct la pinii GPIO de la Raspberry Pi 3.

Utilizați acești pini pentru a conecta LED-uri sau dispozitive similare sau pentru a indica funcționarea releelor.

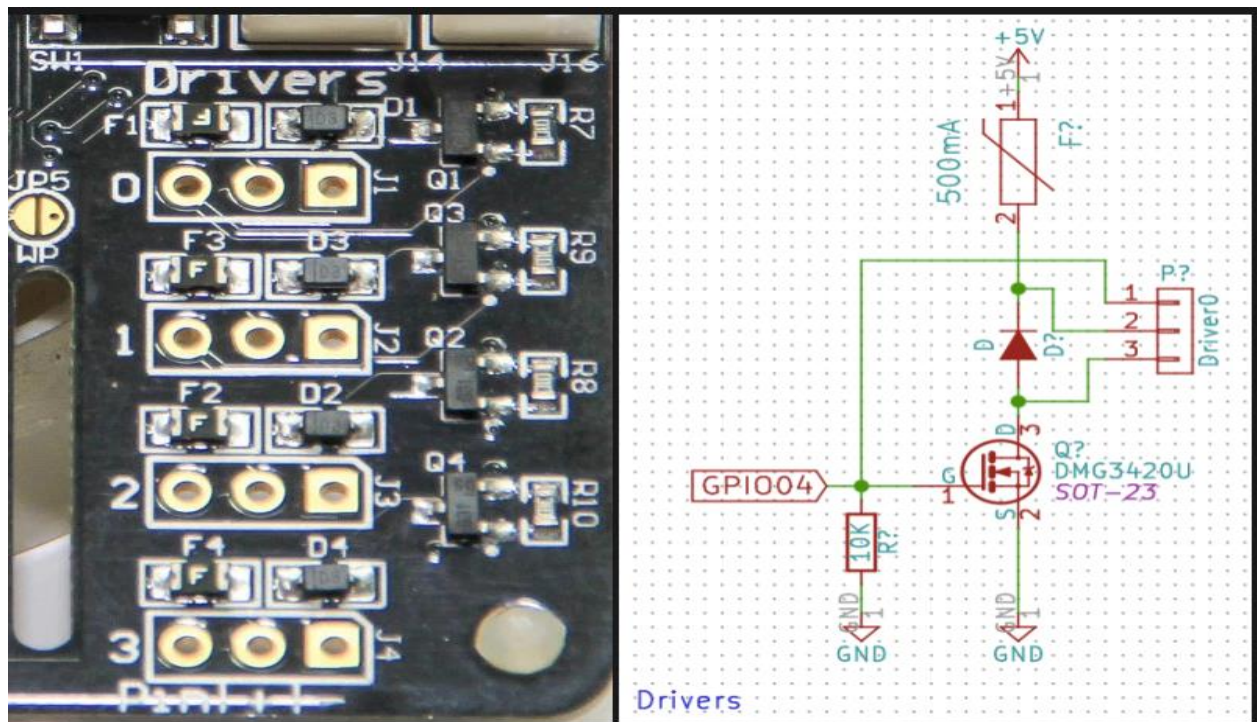


Figura 47: Pini dedicați pentru diverse drivere

Conectori microfon și buton

Lucruri mai interesante se întâmplă în partea dreaptă, unde avem un buton și doi conectori JST. Conectorul cu 4 pini este destinat butonului care se află pe partea superioară a carcasei asamblate. Butonul mic, montat pe PCB, este conectat în paralel cu comutatorul extern și poate fi utilizat în locul său în timpul configurării și testării. JST cu 5 pini este pentru conectorul microfonului . [2] [3]

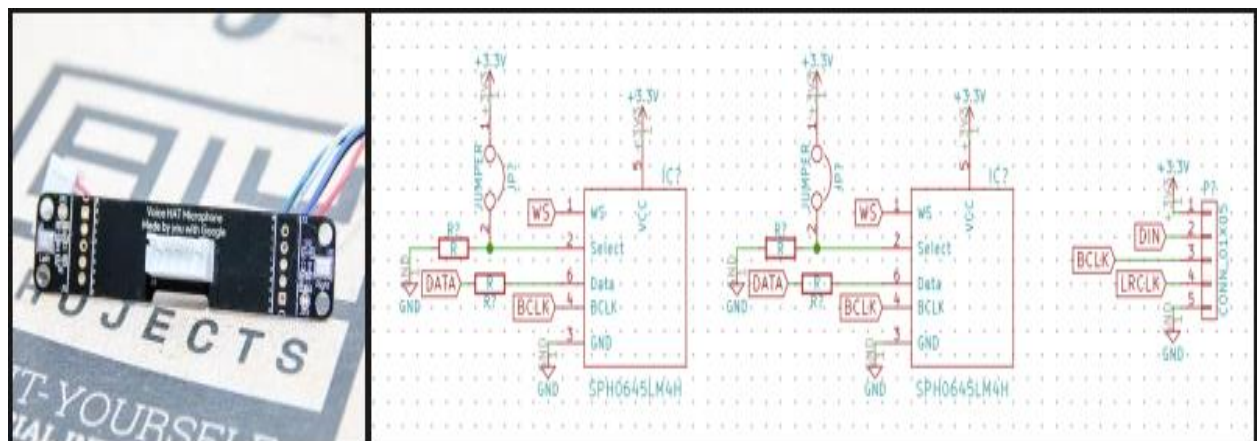


Figura 48: Pini dedicați pentru microfon

CAP.6 Implementarea algoritmilor pentru controlul unor dispozitive hardware

În urma instalării de Amazon Alexa și Google Assistant pe Raspberry Pi, acesta poate recunoaște cuvintele pe care le aude și poate să realizeze căutări pe internet oferind răspunsuri la întrebări, poate să programeze un timer sau poate să oprească Raspberry Pi. Ambii asistenți vocali necesită un declanșator pentru a începe ascultare. La Amazon Alexa vom folosi cuvântul “Alexa”, iar la Google Assistant vom folosi “Ok Google”. Am folosit pentru implementare asistentul de la Google deoarece utilizarea lui a fost mai facilă și am dispus de kit-ul de AiY.

Google Assistant aduce împreună cu instalarea mai multe script-uri pe care le putem folosi în vederea recunoașterii vocale, precum și în vederea răspunsului la cuvintele recunoscute:

- ✓ `assistant_grpc_demo.py` : care așteaptă apăsarea butonului pe care Voice Kit-ul Google îl oferă pentru a începe recunoașterea. După apăsare el doar execută anumite comenzi specificate de utilizator, răspunzând și la întrebări specifice.

if text:

```
if text == 'goodbye':  
    status_ui.status('stopping')  
    print('Bye!')  
    break
```

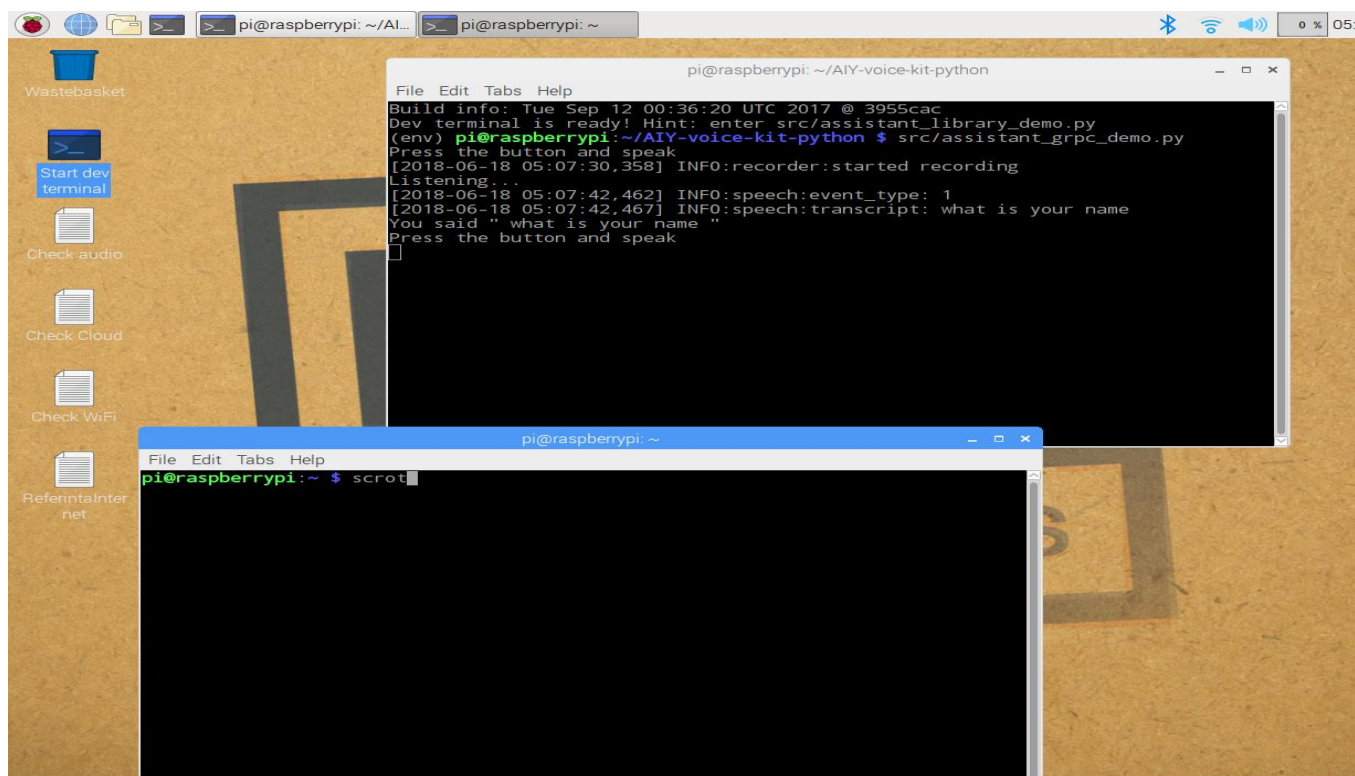


Figura 49: Rularea scriptului și rezultatul recunoașterii în secțiunea “You said”

- ✓ `assistant_library_demo.py` : care așteaptă rostirea cuvintelor “Ok Google” pentru a începe recunoașterea . După ce am spus “Ok Google” asistentul așteaptă de la noi să îi adresăm o întrebare sau să îi oferim o sarcină de îndeplinit de exemplu: “Who is the president of Romania”/”Set a timer for 5 minutes”.Răspunsul la întrebări fiind cautat în baza de date de care google dispune și la care Google Assistant are acces.

```
def process_event(event):
```

```
    status_ui = aiyo.voicehat.get_status_ui()
```

```
    if event.type == EventType.ON_START_FINISHED:
```

```
        status_ui.status('ready')
```

```
        if sys.stdout.isatty():
```

```
            print("Say "OK, Google" then speak, or press Ctrl+C to quit...")
```

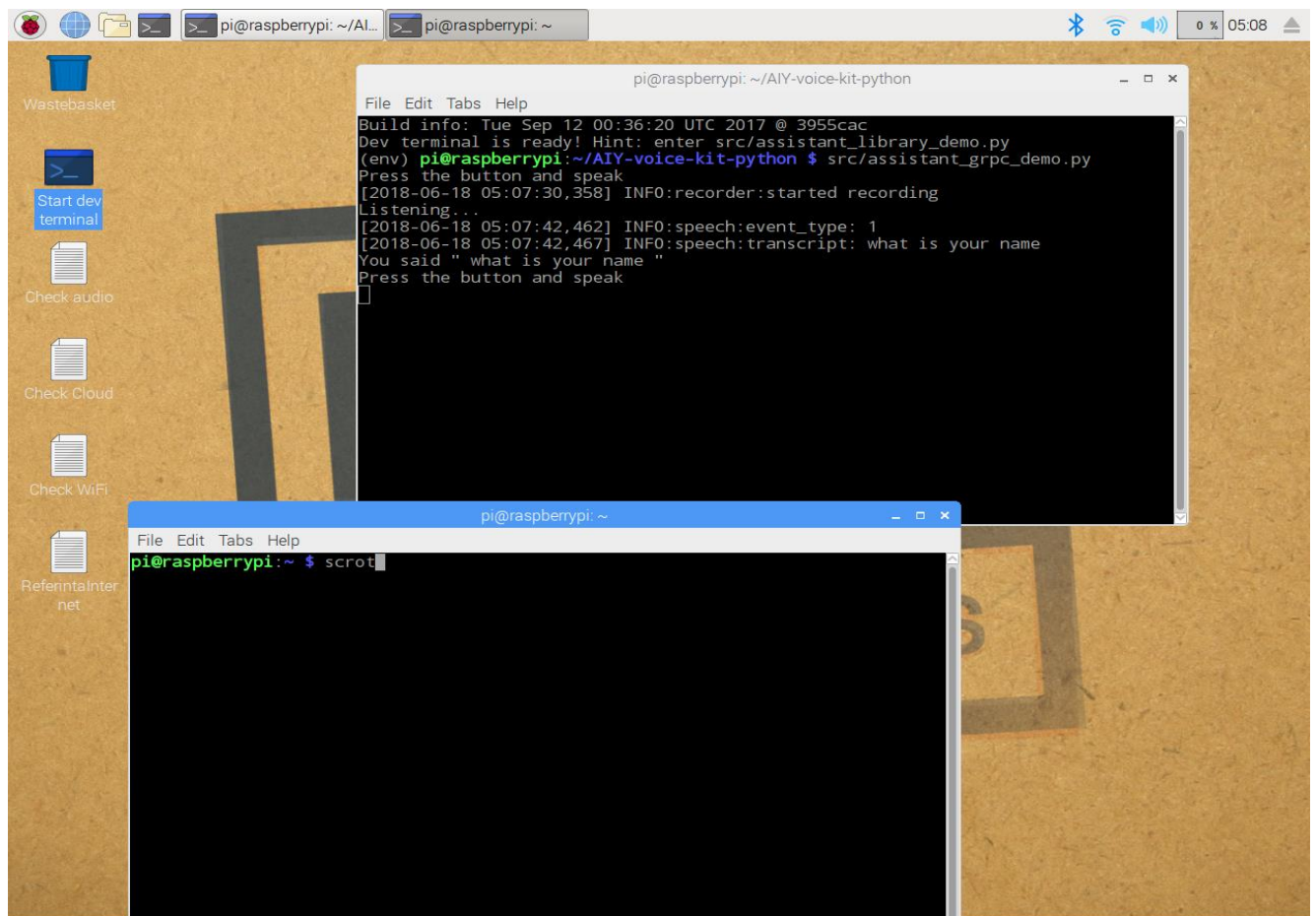


Figura 50:Rularea scriptului pentru recunoașterea comenzii forward

- ✓ `assistant_library_with_local_commands_demo.py` : care așteaptă rostirea cuvintelor “Ok Google” pentru a începe recunoașterea.După recunoaștere el doar execută anumite comenzi specificate de utilizator ,nerăspunzând la întrebări specifice.

```
if text == 'power off':
```

```
    assistant.stop_conversation()
```

```
    power_off_pi()
```

```

elif text == 'reboot':
    assistant.stop_conversation()
    reboot_pi()
elif text == 'ip address':
    assistant.stop_conversation()
    say_ip()

```

- ✓ **cloudspeech_demo.py:** care așteaptă apăsarea butonului pe care Voice Kit-ul Google îl oferă pentru a începe recunoașterea . După apăsare el doar execută anumite comenzi specificate de utilizator ,nerăspunzând la întrebări specifice fiind oferit de google în vederea utilizării la aplicații de control a unor echipamente hardware, dar folosirea lui gratuită este limitată la doar 60 de minute pe luna, depășirea acestui prag fiind taxată cu 0.006\$ pentru 15 secunde.

```

if text is None:
    print('Sorry, I did not hear you.')
else:
    print('You said "', text, '"')
    if 'turn on the light' in text:
        led.set_state(aiy.voicehat.LED.ON)
    elif 'turn off the light' in text:
        led.set_state(aiy.voicehat.LED.OFF)
    elif 'blink' in text:
        led.set_state(aiy.voicehat.LED.BLINK)
    elif 'goodbye' in text:
        os._exit(0)

```

Am ales să creez două scripturi care au la bază recunoașterea folosită în `assistant_grpc_demo.py` și `assistant_library_with_local_commands_demo.py` deoarece acest tip de recunoaștere nu are un timp limitat de folosire ,după expirarea caruia să apară taxe suplimentare și pentru faptul că sunt puse în evidență cele două metode de recunoaștere și anume apăsarea butonului sau rostirea cuvintelor “Ok Google” pentru a declanșa recunoașterea. Am folosit recunoașterea prin apăsarea butonului pentru a controla aprinderea și stingerea unui LED prin rostirea (LED ON și LED OFF). Recunoașterea prin rostirea “Ok Google” pentru a controla un kit de robot cu patru motoare prin rostirea comenzilor (Forward,Backward,Left și Right).

Kit-ul de robot este controlat cu ajutorul unei punți H L298N care este alimentată de o baterie de 12V, iar de la intrările punții IN1, IN2, IN3, IN4 sunt controlați cu pinii de la Raspberry Pi disponibili după aplicarea VOICE HAT-ULUI și anume GPIO (24,26,13,6). Când rostim cuvântul “Forward” mașina trebuie să se deplaseze în față, când rostim “Backward” mașina trebuie să se deplaseze în spate și la fel pentru stânga și dreapta folosind cuvintele “Left” și “Right”.

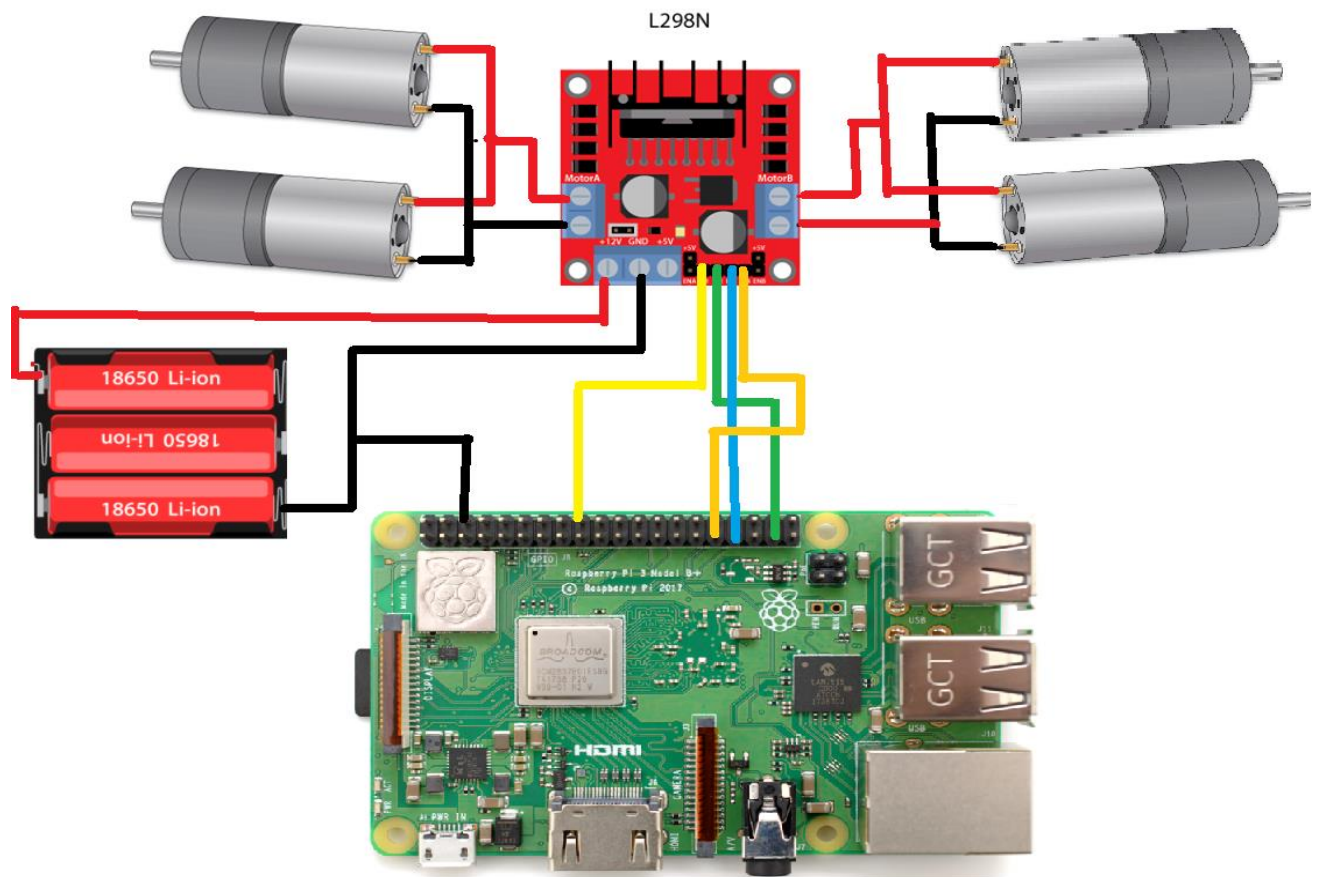


Figura 51: Schema bloc pentru controlul kit-ului de robot

Exemplu de funcție pentru direcția înainte :

```
PWM_FORWARD_LEFT_PIN = 24 # IN1 - Forward Drive
PWM_REVERSE_LEFT_PIN = 26 # IN2 - Reverse Drive
PWM_FORWARD_RIGHT_PIN = 13 # IN1 - Forward Drive
PWM_REVERSE_RIGHT_PIN = 6 # IN2 - Reverse Drive
forwardLeft = PWMOutputDevice(PWM_FORWARD_LEFT_PIN, True, 0, 1000)
reverseLeft = PWMOutputDevice(PWM_REVERSE_LEFT_PIN, True, 0, 1000)
forwardRight = PWMOutputDevice(PWM_FORWARD_RIGHT_PIN, True, 0, 1000)
reverseRight = PWMOutputDevice(PWM_REVERSE_RIGHT_PIN, True, 0, 1000)
frequency = 250
def go_forward():
    forwardLeft.value = 1.0
    reverseLeft.value = 0
    forwardRight.value = 1.0
    reverseRight.value = 0
    time.sleep(4)
    forwardLeft.value = 0
    reverseLeft.value = 0
    forwardRight.value = 0
    reverseRight.value = 0
```

Media recunoașterii corecte 93 % Forward

Media recunoașterii corecte

91 %

Backward

48

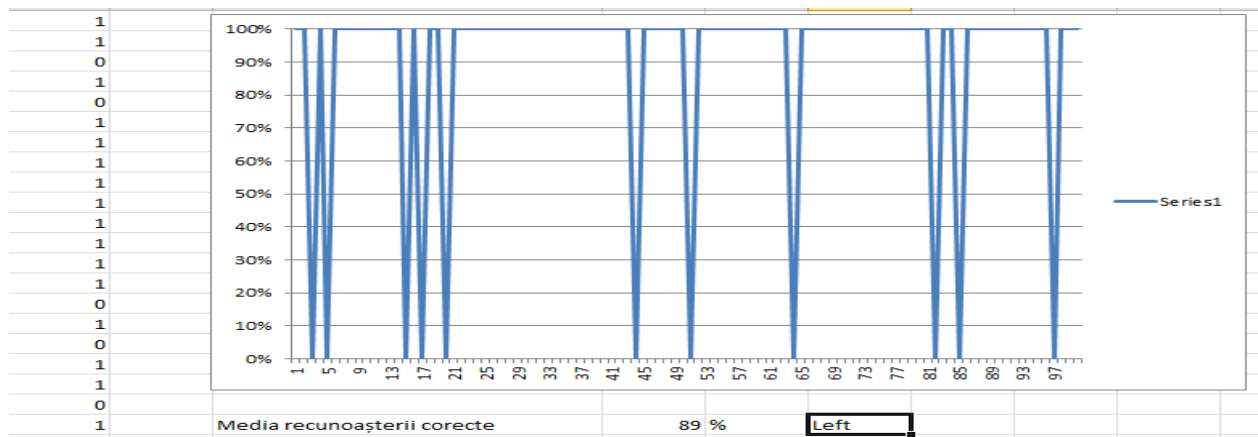


Figura 54: Procentul recunoașterii corecte a cuvântului left pentru 100 de încercări

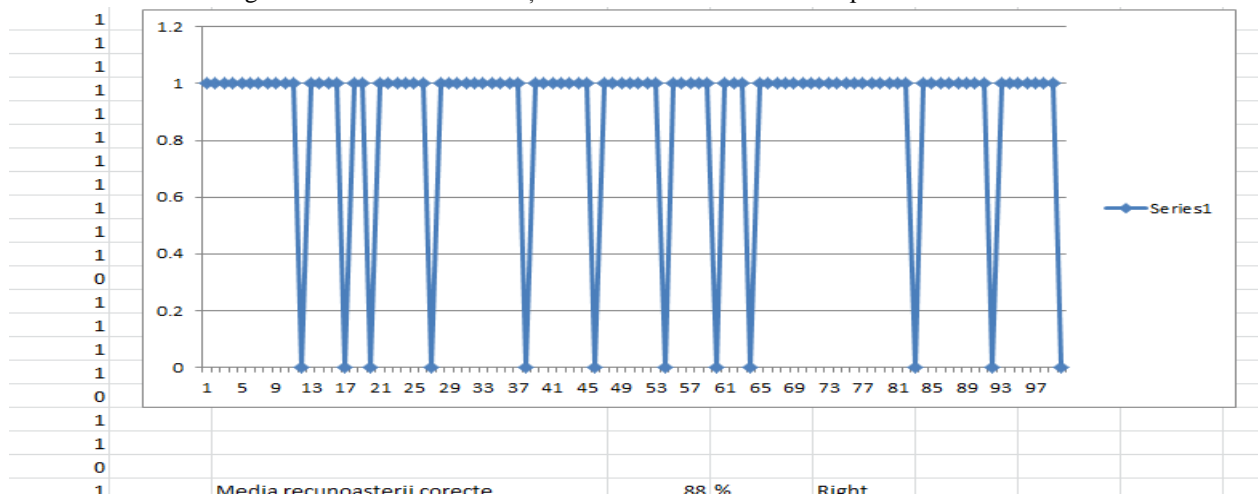
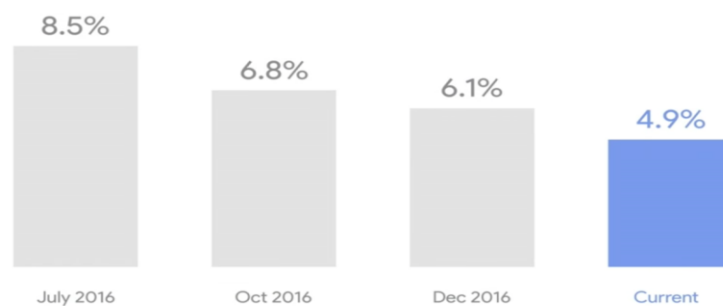


Figura 55: Procentul recunoașterii corecte a cuvântului right pentru 100 de încercări

Speech Recognition

Word Error Rate



US English only.

Figura 55: Procentul oferit de Google pentru acuratețea Google Assistant în 2017

La I / O 2017, Sundar Pichai a remarcat că calculatoarele se îmbunătățesc din ce în ce mai mult în înțelegerea introducerii vocii, iar Google a obținut "progrese semnificative" în recunoașterea vorbirii. De fapt, sistemele Google de învățare a mașinilor sunt acum aproape la egalitate cu oamenii.

Conform Raportului Anual de Tendințe în Internet al lui Mary Meeker, recunoașterea vocală în sprijinul mașinilor Google - începând din mai 2017 - a obținut o rată de precizie de 95% pentru limba engleză. Această rată actuală este, de asemenea, pragul pentru precizia umană.

Cuantificarea progresului Google, acuratețea sa îmbunătățit cu aproape 20% începând cu 2013.

Această realizare este destul de remarcabilă și se aliniază cu Pichai, menționând că "ratele de eroare continuă să se îmbunătățească chiar și în medii zgomotoase".

Eforturile Google în AI ajută la aceste îmbunătățiri. De exemplu, o tehnică deep learning, cunoscută sub numele de formare a rețelelor neuronale, a permis companiei să lanseze Google Home cu numai două microfoane, dar să obțină aceeași calitate ca și opt.

De asemenea, este responsabilă de funcții recente, cum ar fi suportul pentru mai mulți utilizatori, care poate recunoaște până la șase utilizatori diferiți și poate oferi rezultate asistent personalizate.

Rezultatele obținute în urma testării au fost de 93%,91%,89%,88% ,destul de apropiate de datele oferite de google în rapoartele prezentate mai sus.

Din punct de vedere al distanței maxime până la care se poate face o recunoaștere corectă depinde destul de mult de mediul în care se face testarea .

Într-o camera de cămin de 4mX3m am obținut o recunoaștere corectă de la o distanță maximă de 1.8m-2m .

În aer liber rezultatele au fost puțin mai scăzute ,ajungând până la 1.5m-1.8m.

CAP.7 Concluzii

Acest proiect a atins toate punctele propuse și poate constitui baza unui proiect mai mare ce include și elemente avansate de inteligență artificială pentru face această temă și mai complexă.

Domeniul recunoașterii vocale bazat pe inteligență artificială și algoritmi de deep learning este unul destul de nou ,rezultatele semnificative au venit spre sfârșitul anului 2016 ,dar care merită cu siguranță studiat mult mai profund în viitor, având exemple precum (Siri , Google Assistant și Amazon Alexa)

Posibilitățile de automatizare oferite de domeniul recunoașterii vocale sunt foarte multe și diverse ,putând să fie de ajutor mai ales pentru oamenii cu probleme de vedere sau probleme locomotorii în vederea ușurării vieții acestora ,dar și pentru oamenii normali care își pot face viața mai ușoară și mai rapidă ,câștigând mult timp .

Performanțele robotului realizat și a asistenților utilizați sunt foarte dependente de locul unde se află, pentru o recunoaștere cât mai corectă a comenzilor ,dar și de vocea celui care îl controlează,desigur și de distanța de la care este controlat ,aceasta fiind limitată în jurul unei distanțe maxime de aproximativ 2 metri.

Folosirea platformei Raspberry Pi pentru îmbunătățirea acestui proiect în viitor este esențială,menținând un cost redus și permițând folosire unei multitudini de limbaje de programare și având o viteză de execuție foarte bună .

De asemenea oferă avantajul rulării unor distribuții de linux ca sistem de operare , lucru foarte important având în vedere apropierea acestui sistem de operare de platforme pe Android sau IOS.

Dacă problemele de recunoaștere și de distanță vor fi soluționate,iar inteligența artificială va beneficia de aceeași dezvoltare fulminantă ca în ultimii ani , eu cred că putem deveni martori la o schimbare radicală a lumii în care trăim acum din toate punctele de vedere.

Referințe bibliografice

www

- [1] Revista MagPi din martie 2017 Issue 55 paginile 42-47 :
<https://www.raspberrypi.org/magpi-issues/MagPi55.pdf>
- [2] Revista MagPi din mai 2017 Issue 57 paginile 15-31:
<https://www.raspberrypi.org/magpi-issues/MagPi57.pdf>
- [3] Revista MagPi din iunie 2017 Issue 58 paginile 66-73:
<https://www.raspberrypi.org/magpi-issues/MagPi58.pdf>
- [4] <https://aiyprojects.withgoogle.com/voice>
- [5] <https://lifehacker.com/how-to-build-your-own-amazon-echo-with-a-raspberry-pi-1787726931>
- [6] Wikipedia Bixby :
[https://en.wikipedia.org/wiki/Bixby_\(virtual_assistant\)](https://en.wikipedia.org/wiki/Bixby_(virtual_assistant))
- [7] Wikipedia Siri :
<https://en.wikipedia.org/wiki/Siri>
- [8] Wikipedia Google Assistant :
https://en.wikipedia.org/wiki/Google_Assistant
- [9] Wikipedia Amazon Alexa :
https://en.wikipedia.org/wiki/Amazon_Alexa
- [10] Roborace.com

Cărți

- [11] Ivan Bogdanov, Conducerea Roboților , Editura Orizonturi Universitare, ISBN 978-973-638419-6