

Beszéd felismerés könnyen PCLinuxOS-on

PCLinuxOS Magazine – 2018. szeptember

Írta: Alessandro Ebersol (Agent Smith)



A beszéd felismerés a számítástechnika egyik legérdekesebb aspektusa. Hatalmas előrelépések történtek az elmúlt években. Ezek a fejlesztések átkerültek a személyi számítástechnikába is. Így, a hang felismerés technológiája területén elért eredmények elérték a otthoni felhasználót.

A beszéd felismerés olyan tulajdonság, ami a hozzáférésben, vagy hosszú szövegek begépelésében elfáradva alkalmazható. Igen, hosszú szöveg bevitele során megfáradt kezeid pihentetése különösen kellemes.

Más operációs rendszerek használói egyszerűen elérhetik a beszéd felismerést: Windowsban van a Dragon Naturally Speaking, Macben a Dictanote és a WordQ SpeakQ. Mi a helyzet a Linuxszal?

A Linuxban számos beszéd felismerési eljárás van. Ugyanakkor, nagyon komplikált módon lehet azokat munkára fogni. Általában ezek pocketsphinx-et, egy beszéd felismerő és szöveg-beszéd motort igényelő megoldások.

De nem kell félni, mivel a hang felismerés a Google-nek hála (most már Alphabet Company) lassan megérkezett és a PCLinuxOS-felhasználók

kihasználhatják a legnagyobb internetes keresőmotor nyújtotta előnyöket.

A Google Voice Voice Recognition Infrastructure, Google cloud voice API

A Google kiadta az új Cloud Speech API-ját – a San Francisco-i NEXT-en – egy korlátozott előzetes változatát fejlesztők számára 2016-ban. Ezt a beszéd felismerő technológiát fejlesztették tovább és már használják egy ideje különféle Google-termékekben, mint a Google keresőmotorja, ahol lehetőség van hangalapú keresésre.

A képesség, ami a beszédet szövegre konvertálja, összetett idegpálya-hálózatokon alapul, illetve legmodernebb öntanító gépi algoritmusok bizonyultak különösen hatásosnak a hang- és videó jelekben, formulák felismerésében. A neurális hálózat frissült, mivel a Google új beszéd mintákat gyűjtött, így új kifejezéseket tanult, és folyamatosan növekszik a felismerés pontossága. Ez a technológia működteti az OK Google-t, az Android-ot, a Google Chrome-ot és a Google Virtual Assistant-ot. A felhőalapú beszéd felismerő technológiájával a Google a régi Windows Dragon Dictate-et (most Dragon Natural Speaking) készítő Nuance-szal kelt versenyre.

Beszéd felismerés felhőben

A beszéd szövegre konvertálás opciót olyan esetekben használják, mint hanggal vezérelt varázslók a mobil eszközökben, lakásautomatizálás, beszéd leiratok és telefonhívások automatikus titkosítása.

Most ez a technológia elérhető lesz egy fejlesztői felhő-szolgáltatás formájában, lehetővé téve bármely alkalmazás számára, hogy hang szöveggé felismerést integráljon, így adva használható alternatívát az elterjedt Nuance (például az Apple Siri-ben és a Samsung S-Voice-ban alkalmazott) technológiájával szemben, illetve versenyezve más megoldásokkal, mint az IBM Watson Speech to Text és a Microsoft Bing Speech API.

Most hogyan profitáljunk a Google Voice Cloud Recognition-ból?

Élvezhetjük a Google által a termékeibe integrált előnyöket, mivel számos oldal kínál hang felismerési szolgáltatást, a Google cloud voice (felhő hang) API-ján keresztül.

Ezen szolgáltatások előnyeinek kiaknázására Internet-kapcsolat, számítógéphez kapcsolt mikrofon és Google Chrome böngésző használata kell. A továbbiakban néhány ilyen szolgáltatást vizsgáló meg.



<https://speechnotes.co>

A Speechnotes nagy teljesítményű, online jegyzetfüzet tiszta és hatásos dizájnnal elképzelésid megvalósítását támogatóan kialakítva, így a gondolataidra koncentrálhatsz. A honlap működtetői a legjobb diktáló eszközre törekszenek, felhasználva a legfejlettebb beszéd felismerő technológiát, ami napjainkban csak elérhető csakúgy, mint integrálva

belső (automatikus, vagy kézi) hatékonyságot, teljesítményt és kényelmet növelő eszközöket. Teljesen online működésű a Chrome böngészőben. Letöltés, telepítés illetve regisztrálás sem szükséges, azonnal elkezdheted dolgozni.

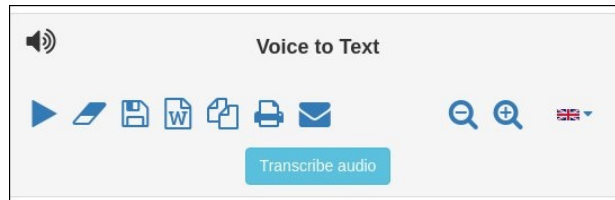
Értékelésem: támogatja a brazil portugált és több mint 40 nyelvet, Google Drive-ra ment és ha regisztrálsz a cég szolgáltatására, a dokumentumaidat bármely, a fiókodba bejelentkezett gépről elérheted. Én használtam. A beszéd felismerése gyors és pontos. Google Chrome-on működik és mobil telefonnal is.



<https://www.spechtexter.com/>

A SpeechTexter egy professzionális online szövegkonvertáló arra tervezve, hogy egyszerűsítse és felgyorsítsa a munkádat. A cég a legjobb online átíró gyakorlat megvalósítására törekszik innovatív beszéd felismerési technológiával és nagy pontossággal. Ezt a technológiát csak a Chrome böngésző támogatja. Más böngészők még nem alkalmazzák a beszéd felismerést. Nincs regisztráció, vagyis azonnal dolgozhatsz.

Értékelésem: támogatja a brazil portugált és több mint 40 nyelvet. TXT-be ment, nyomtat, vagy vágólapra másol. Nincs regisztráció, ami annyit tesz, hogy csak egy gépen használható egy időben. A teszt során kicsit lassabb volt mint a Speechnotes.



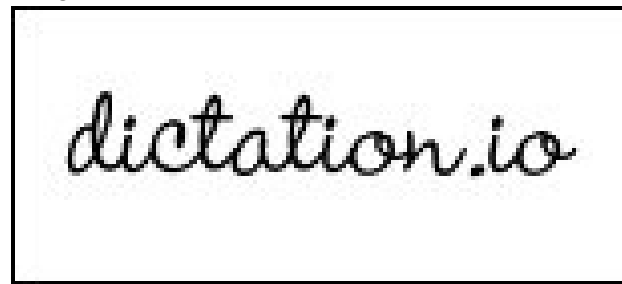
<https://www.textfromspeech.com>

Ez egy beszéd szövegre, szöveget beszédre szolgáltatás. Mit jelent? Azt, hogy egyszerűen diktálhatsz és egyetlen ujjad megmozdítása nélkül írja le. Más részről, gépelhetsz szöveget és a honlap által nyújtott szolgáltatás hangosan felolvassa.

A következő jellemzői vannak:

- ingyenes és online;
- nincs letöltés, telepítés, vagy regisztrálás. Multiplatformos;
- csak 7 nyelvet támogat (a brazil portugál nincs közte);
- a diktálás szüneteltethető és megállítható (megjegyzi az utolsó szó helyét);
- felismeri a kimondott írásjeleket: pl. „Comma” (vessző) kimondására „, ”-t gépel;
- okos nagybetűsítés;
- menthető, másolható, nyomtatható, vagy továbbítható a szöveg.

Értékelésem: mivel a brazil portugált nem támogatja, nem próbáltam ki. Úgy tűnik, hogy ez rendelkezik a legkevesebb erőforrással a vizsgált szolgáltatások közül.



<https://dictation.io/>

A dictation.io felismeri és leírja az elterjedtebb nyelveket, közte angolt, spanyolt, franciát, olasz, portugált és másokat. A támogatott nyelvek teljes listáját itt nézheted meg: <https://dictation.io/languages>. Új bekezdések, írásjelek, smiley-k és más speciális karakterek adhatók hozzá egyszerű szóbeli parancsokkal. Például a „new line” új sorba viszi a kurzort, vagy a „Smiley face” viszi be a „ :-) ”-t. A támogatott parancsokról külön lista van.

Értékelésem: lehetővé teszi számos művelet végrehajtását a felismert szöveg alapján. Vágólapra másol, ment TXT-be, elküldi Tweet-ként, vagy e-mailben, nyomtatható, illetve exportálható PDF-ben, vagy a korábban számítógépbe beolvasott szöveget a gép diktálhatja. Sebessége a Speechnotes-hoz hasonló.

A négy megvizsgálatnál sokkal több lehetőség van, de az elemzésük több kutatási időt igényelne az elemzések miatt. Ami számít, hogy amennyiben hosszú szövegeket kell diktálnod a PCLinuxOS tud segíteni a végrehajtásban. Ami a beszéd felismerést illeti a Linux már nem másodosztályú állampolgár többé.

Remélem tetszett. Ezt a szöveget a Speechnotes segítségével készítettem el és LibreOffice-szal szerkesztettem meg.

Élvezd ki az új lehetőségeket és találkozzunk a jövő hónapban! Anya nézd! Kéz nélkül írtam le!

