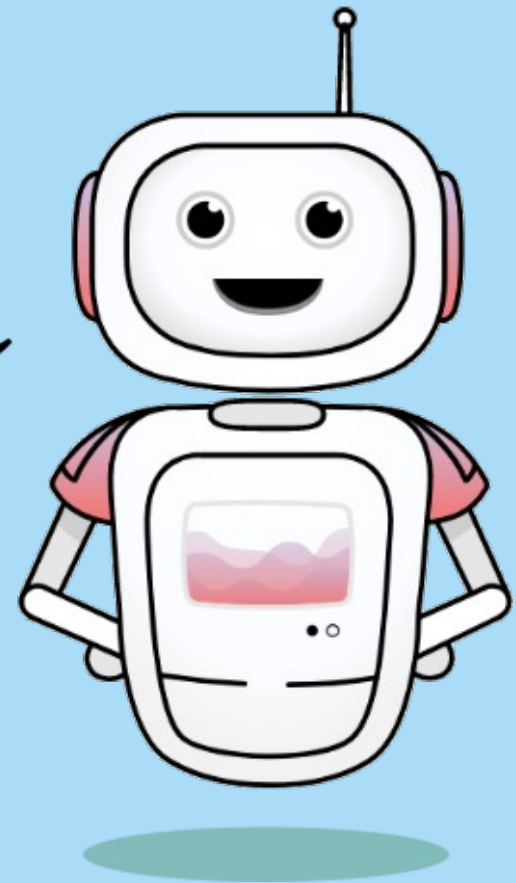


**Zer da  
Common Voice?**



# Mozilla fundazioaren egitasmoa

Produktuak:

## 1 Datuak

Audio grabaketen datu-sorta librea.

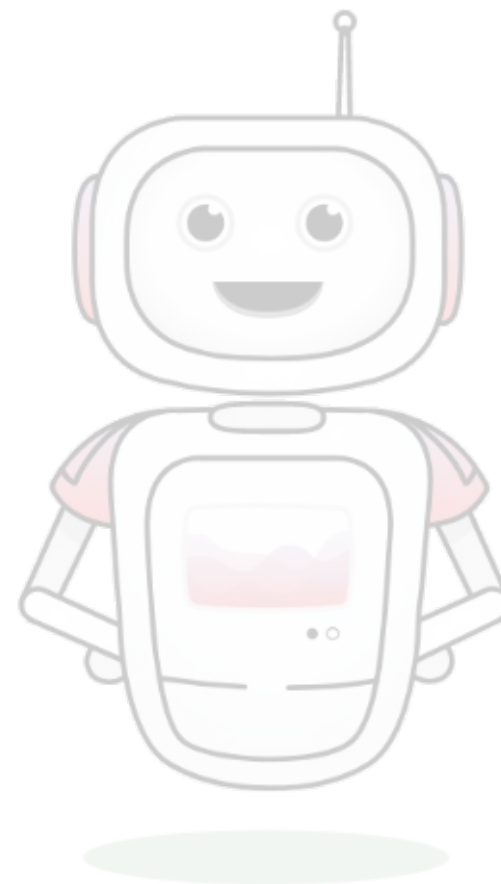
## 2 Softwarea

Kode irekiko ahots ezagutzailea (STT motorra).

Hizkuntzak:

54 hizkuntza martxan (euskara barne).

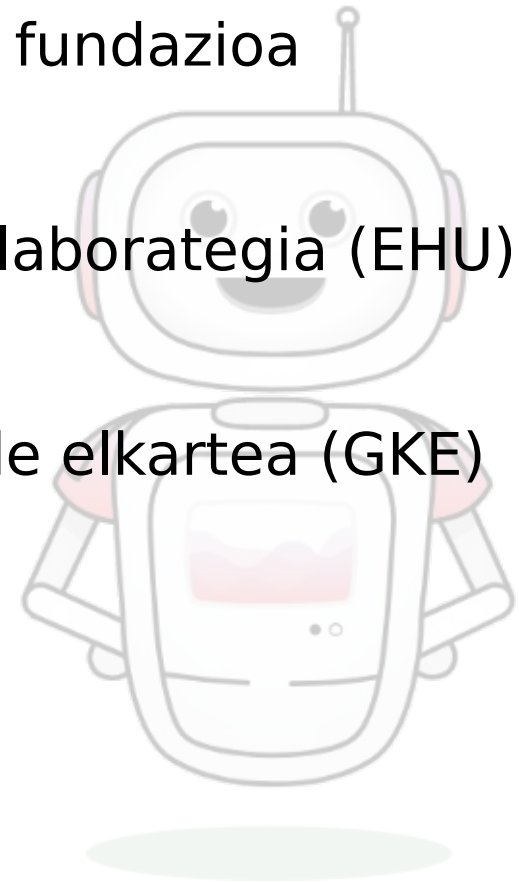
68 hizkuntza bidean.



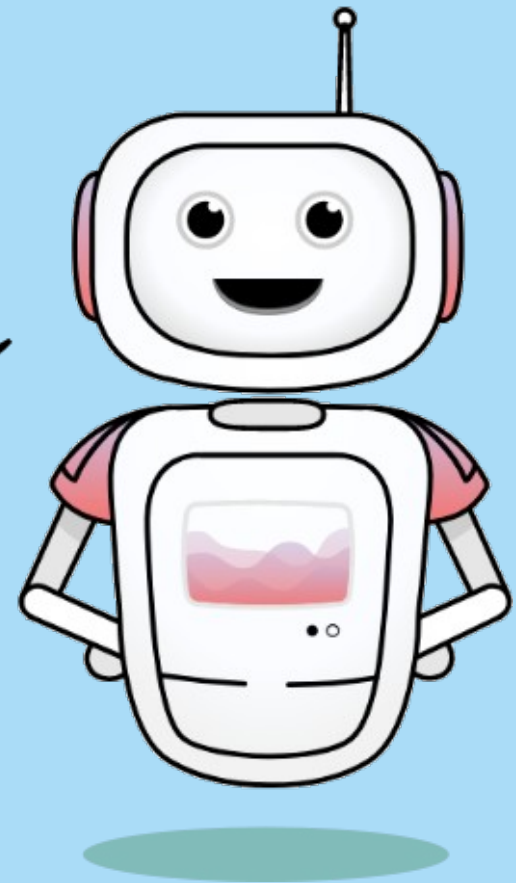
## Partaideak (euskara atalean)

- **1.000 hiztun boluntario**
- Librezale taldea
- Wikipedia fundazioa
- Argia aldizkaria
- iametza enpresa

- Ixa taldea (EHU)
- Elhuyar fundazioa
- Aholab laborategia (EHU)
- Garabide elkarte (GKE)



**Zergatik da  
garrantzitsua?**



# Egoera

- Ahotsa gure artean komunikatzeko modurik naturalena da.
- Ahotsaren ezagutza gure gailuetara heldu da.
- Ahots-datuak garestiak eta jabedunak dira.
- Euskaraz egokiak oso gutxi daude!
- Ikasketa automatikoan (*machine learning*) datu andana behar da.
- Euskara ez da errentagarria sektoreko enpresentzat.
- Euskaldunak makinekin erdaraz ari dira!

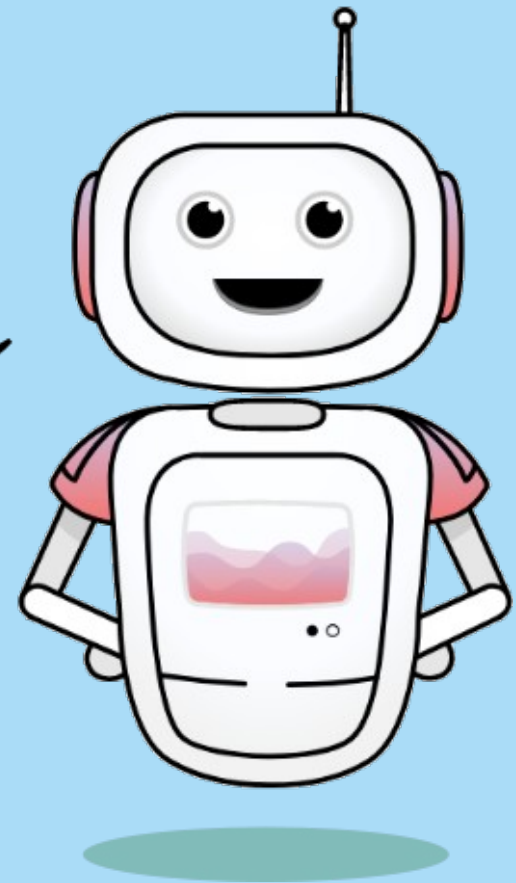


# Helburuak

- Aplikazio txundigarriak ahaldundu:
  - Denbora errealeko itzultzaileak.
  - Ahots bidezko asistenteak.
  - ...
- Sektorea indartu eta demokratizatu.
- Inklusibitatea handitu:
  - Maiz adin ertaineko gizonezkoek eginak.
  - Maiz emakumezkoen ahotsa ezagutzeko zailtasunak.
  - Adin, sexu eta azentu desberdinak sustatu.



**Nola lagundu?**



## 1 **Esaldiak biltzen**

Euskarazko testuak biltzen: itzulpen memoriak...

## 2 **Grabaketak egiten (zeharka)**

Hutsune estatistikoak betetzen, eragileak erakartzen...

## 3 **Grabaketak balioztatzen**

## 4 **Ideiarik? Proposamenik?**

