

Euskarazko hizkuntza teknologia libreak

Lekua

Euskarabildua 2018

San Telmo Museoa, Donostia, Gipuzkoa

euskarabildua.eus

Hizlaria

Marko Txopitea

ikusimakusi.eus

Adimen artifizialaren helburuak:

Pentsatzea/hausnartzea

Planifikatzea

Automatikoki ikastea (*machine learning*)

Hizkuntza naturala prozesatzea (NLP)

Ikusmen artifiziala

Robotika

Adimen artifizial orokorra

Mekanismoak:

Hitz-prozesamendua

Irudi-prozesamendua

Ahots-prozesamendua

Mekanismoak eta teknikak:

Hitz-prozesamendua

- Lexikoa, gramatika, semantika
- Itzulpen automatikoa
- Polaritate detekzioa
- Fonetika (ahotserako)

Irudi-prozesamendua

- Karaktere ezagutza optikoa (OCR)
- Hizkien entrenamendua
- 3D avatarrak: lip sync (ahotserako)

Ahots-prozesamendua

- Sintesia: testutik ahotsera (TTS)
- Ezagutza: ASR/STT
- Dialektoak, generoak, adinak...
- Hizlari -baliozkotzea/-identifikazioa

Agente garrantzitsuenak:



Hitz-prozesamendua

Irudi-prozesamendua

Ahots-prozesamendua



Euskarazko hizkuntza teknologia libre gehienak non aurkitu:

Elhuyar fundazioa:

<https://github.com/Elhuyar>

<http://komunitatea.elhuyar.eus/ig/hizkuntza-baliabideak-eta-tresnak/>

Ixa (EHU):

<https://github.com/ixa-ehu>

<http://ixa.si.ehu.es/produktuak>

Aholab (EHU):

<https://sourceforge.net/u/aholab/profile/>

<http://aholab.ehu.eus/aholab/products/>

Eusko Jaurlaritza:

euskara.euskadi.eus - Euskararen IKTen inbentarioa

Euskarazko hizkuntza teknologia libre esanguratsuenak:



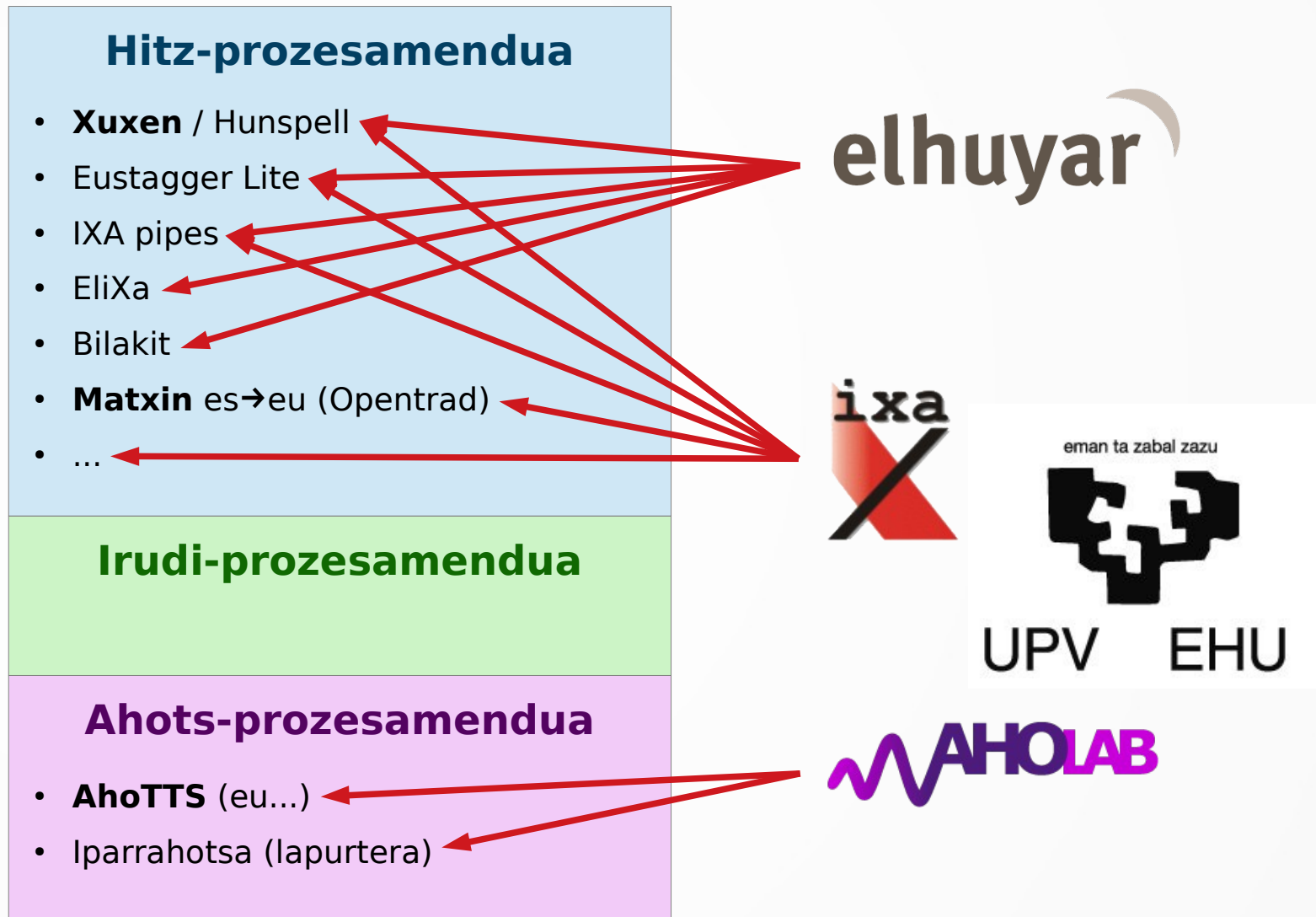
Hitz-prozesamendua

- **Xuxen** / Hunspell
- Eustagger Lite
- IXA pipes
- EliXa
- Bilakit
- **Matxin** es→eu (Opentrad)
- ...

Irudi-prozesamendua

Ahots-prozesamendua

- **AhoTTS** (eu...)
- Iparrahotsa (lapurtera)



Euskarazko beste proiektu aipagarri batzuk:

1) Euskalterm



2) AnySoftKeyboard



3) Mitzuli



4) Common Voice

moz://a




ARIKETA: asmatu hauetako zein EZ den librea!

AnySoftKeyboard Euskaraz

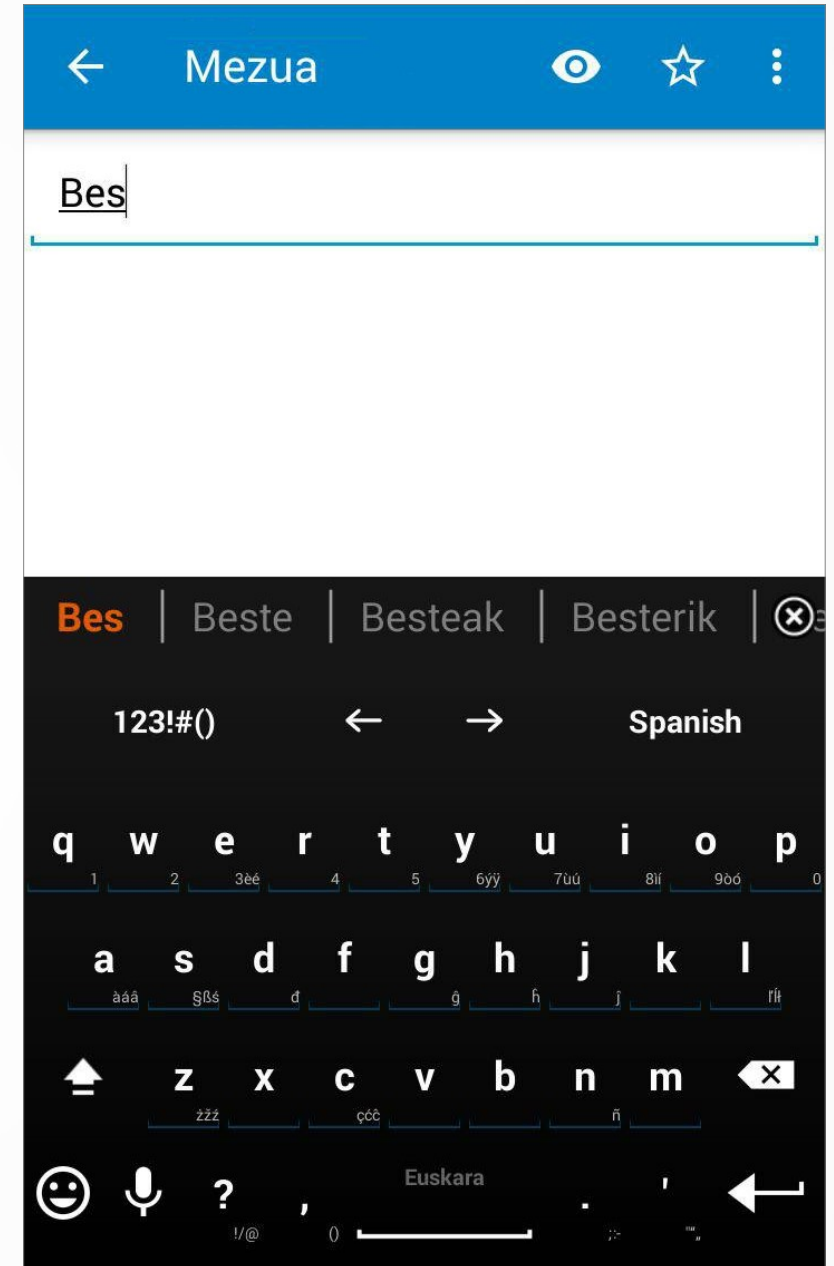
Android OSrako teklatu librea

40 hizkuntzatan erabilgarri.
Hitzak proposatzen ditu.
Laster, keinuen bidezko idazketa ere.

Nortzuk:

-  Asier Iturralde, Iametza
-  Elhuyar Fundazioa (Igor Leturia)
-  Menny Even-Danan, AnySoftKeyboard

 anysoftkeyboard.github.io
 Artikulua Librezeleko blogean



Mitzuli

Android OSrako itzultzaile libre

- Itzulpen automatikoa
- 50 hizkuntza bikote: eu→es, es→eu...
- Kamera bidezko itzulpena (OCR)
- Ahots bidezko sarrera (ASR)
- Ahots bidezko irteera (TTS)

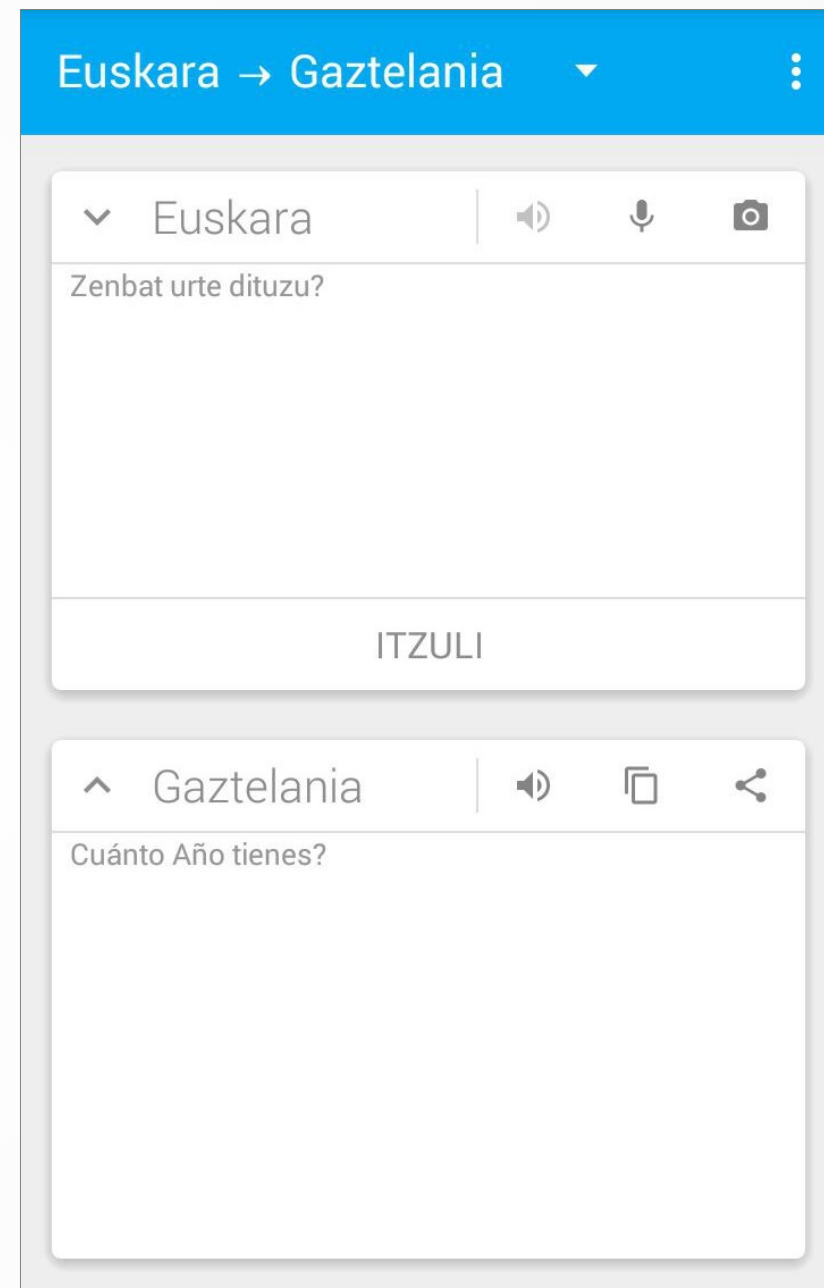
Nork:

 Mikel Artetxe

 Tresnak: Matxin, Apertium, Tesseract-ocr...

 www.mitzuli.com

 [Mikel Artetxeren aurkezpena \(12 minutu\)](#)



Common Voice

STT motor libre

Off-line ibiliko da: segurtasuna, pribatutasuna.
73 hizkuntza bidean, euskara barne.

Faseak:

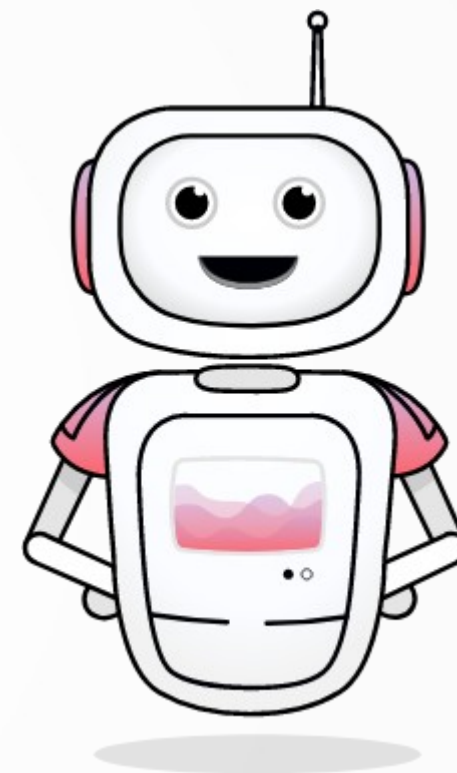
- 1 Prestaketak: euskarazko CC0 testurik?
- 2 Grabaketak
- 3 Pipsqueak motorra

Nortzuk:

- 👤 Librezale taldea
- 👤 Mozilla fundazioa

Tresnak:

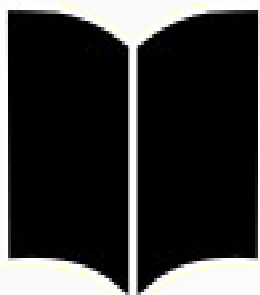
- ⚙️ DeepSpeech
- ⚙️ TensorFlow
- ⚙️ Web Speech API



🔗 voice.mozilla.org

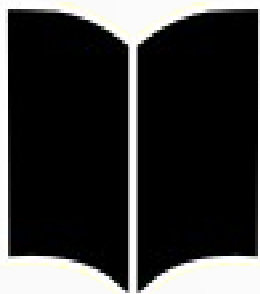
🔗 <https://librezale.eus/wiki/EdukiakJabetzaPublikoan>

Liburu gomendagarriak:



Inmaculada Hernáez, Eva Navas, Igor Odriozola, Kepa Sarasola, Arantza Diaz de Ilarraza, Igor Leturia, Araceli Diaz de Lezana, Beñat Oihartzabal, Jasone Salaberria. **Language Technology Support for Basque**. 2012, pp.48-65. In: Rehm G., Uszkoreit H. (eds) The Basque Language in the Digital Age. White Paper Series. Springer, Berlin, Heidelberg

 https://link.springer.com/chapter/10.1007/978-3-642-30796-6_9



Itziar Aduriz, Iñaki Alegria, Xabier Artola, Arantza Díaz De Ilarraza, Kepa Sarasola. **Teknologia garatzeko estrategiak baliabide urriko hizkuntzetarako: euskararen eta Ixa taldearen adibidea**. Linguamatica, 2011, 3 (1), pp.13-31. <artxibo-00612912>

 https://www.researchgate.net/publication/277045199_Teknologia_garatzeko_estrategiak_baliabide_urriko_hizkuntzetarako_euskararen_eta_Ixa_taldearen_adibidea

Eskerrik asko! Galderarik?

Oharrak:



Aurkezpen honetako testuak, taulak eta diagramak, legeak ahalbidetzen duen neurrian, jabetza publikoan uzten dira. Lizentzia: [CC0 1.0](#).



Erakunde desberdinen logotipoen jabetza eta egile eskubideak berauei dagozkie.

11. orriko robotaren irudiak [CC BY-SA 4.0](#) lizentzia du eta Mozilla fundazioak egin du.



Deskargatu ezazu dokumentu hau [ikusimakusi.eus](#) webgunetik [ODP](#) eta [PDF](#) formatuetan.