

hizkuntza

Language

Hizkuntzaren teknologia, adimen artifizialaren hurren go jomuga

Urteko Txostena
2021

SARRERA



words

language

2

hizkuntza

Hizkuntza ulertzea da adimen artifizialaren helburu nagusietako bat, eta nahiz eta ikasketa sakonari esker aurrerabide handia egin den, adimen artifizialaren ahulgune nagusietakoa izaten jarraitzen du. Izan ere, ordenagailuak ez dira adimentsuak izango harik eta hizkuntza ulertzen duten arte. Transformazio digitalaren eta adimen artifizialaren erronka eta aukera berrien aurrean, ezinbesteko bihurtu da hizkuntza teknologien ikerketan sakontzea. Horretarako sortu genuen HiTZ, helburu bikoitz batekin: hizkuntza teknologien arloan nazioarteko erreferentea izatea, eta Euskararen tratamendu konputazionalan erreferentziazko zentro bihurtzea. Horretarako, zentroak bost ildotan du jarduera: prestakuntza, ikerketa, berrikuntza eta transferentzia, gure

center

```

def assign_position
  # ...
end

def assign_position
  # ...
end

def assign_position
  # ...
end

```

```

def assign_position
  # ...
end

def assign_position
  # ...
end

def assign_position
  # ...
end

```

```

RESULT_TYPES = %w[usual]

```

```

def members
  team1.members ++ team2
end

```

```

end

```

```

def assign_position
  # ...
end

```

kultura bultzatzea, eta gizartean eragitea. Eremu zabal horrek disziplina arteko talde aniztuna eskatzen du, besteak beste, informatikariak, hizkuntzalariak, ingeniariak, itzultzaileak eta soziologoak biltzen dituena. Hitz zentroa osatzen duten Aholab eta Ixa taldeetako ikertzaileen lanak, fruituak eman ditu aurtengoan ere. Azpimarratzekoak dira euskararako lehen hizkuntza eredu neuronal arrakastatsua sortu izana, 2025erako hizkuntza teknologiarako estrategia definituko duen ELE proiektu europarraren lidergoan egotea, CLARIN ikerkuntza-azpiegitura europarrean emandako aurrerapausoak, eta nazioarteko kongresu hoberenetan eta hainbat lehiaketa zientifikotan (COVIDaren inguruan Estatu Batuetako gobernuak bultzatu zuena barne) jasotako 5 sariak. Bereziki aipagarriak izan dira

jaso ditugun Informatikako bi Ikerketa Sari Nazionalak eta gure zuzendariak jasotako errekonozimendua, gure arloko ikerketa-elkarte nagusiak Europan dituen 15 kideen artean hautatua izan baita. Azkenik, gure doktore batek Europako Adimen Artifizialaren inguruko tesi onenaren saria (EurAI) jaso duela ere aipatzekoa da. Oinarri sendo horrek indarberrituta jarraitzeko aukera ematen digu, eta era berean, bestelako zentro teknologiko, enpresa eta erakundeekin elkarlanean, Euskadi hizkuntza-teknologiaren eta adimen artifizialaren nazioarteko erdigune bihurtzeko abagunea eskaintzen digu.

Eneko Agirre (HiTZeko zuzendaria) eta German Rigau (HiTZeko zuzendariordea)

words

words

hizkuntza

HITZ ZENBAKITAN

Ikerkuntza eta transferentzia

34

proiektu

22

transferentzia-
kontratu

65

argitalpen








4

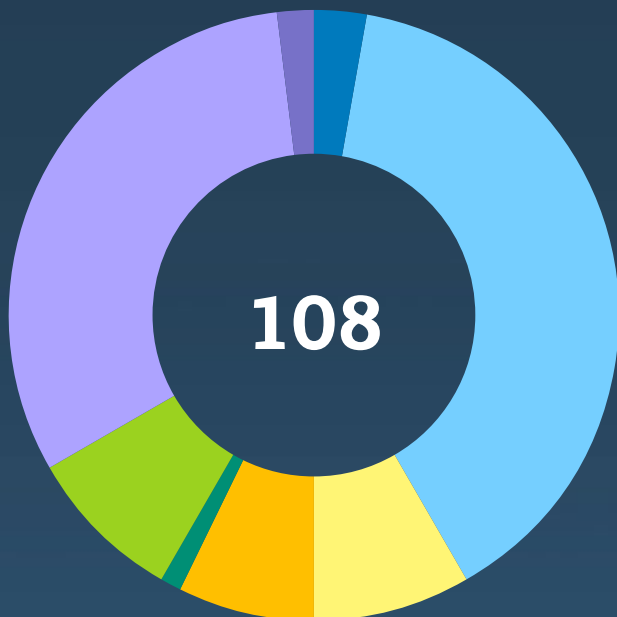
HIZKUNTZA TEKNOLOGIAKO EUSKAL ZENTROA

techn

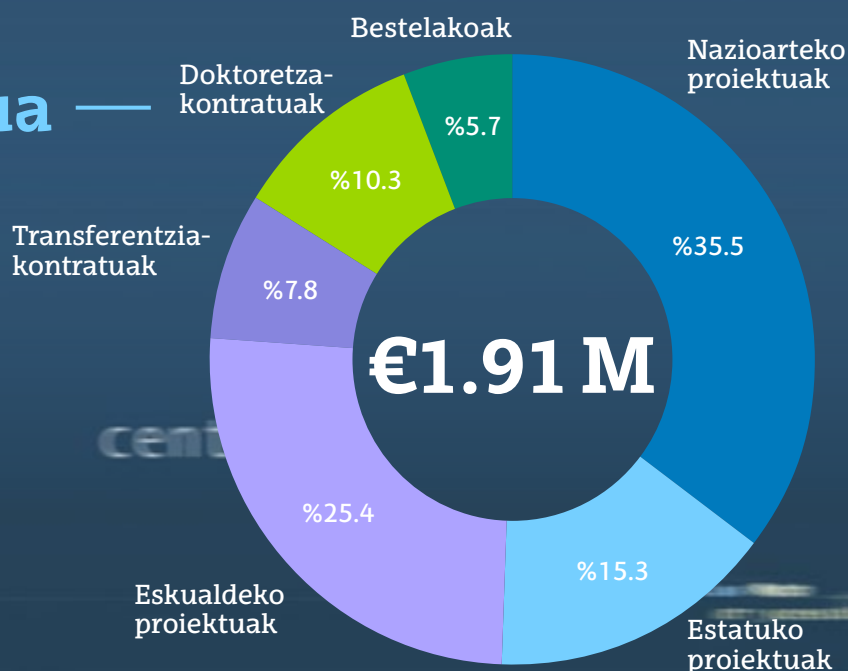


Kideak — 108

| | | |
|--|--------------------------------|----|
|  | Kudeaketa eta administrazioa | 3 |
|  | Irakasleak | 42 |
|  | Doktoretza-bekak | 9 |
|  | Doktoratu ondoko ikertzaileak | 8 |
|  | Ramón y Cajal | 1 |
|  | HAP/LAP ikasleak | 9 |
|  | Doktoretza-ikasleak | 34 |
|  | Doktoretza-tesiak defendaturik | 2 |



Aurrekontua —



hitzak

language

cent

ANTOLAKUNTZA

Hizkuntza ardatz duen adimen artifizialean diharduen ikerketa-zentroa da HiTZ. Euskal Herriko Unibertsitateko zazpi sailetakoko kideek osatzen dutelarik, diziplina anitzeko taldea da. Hizkuntzaren eta hizketaren teknologiak ikertzea da zentroaren helburua, enpresei ezagutza eta teknologia transferitzeko ahalegin nabarmena eginez. Bi ikerketa taldek osatzen dute: **Aholabek** eta **Ixak**. Bi taldeok esperientzia handia dute, 1993az geroztik oinarrizko ikerketa egin, hizkuntza-baliabideak eta -tresnak sortu eta hainbat produktu merkaturatu dituztela. **CLARIN-K Center**-eko kide sortzaileetako bat da HiTZ zentroa, **CLAIRE**-ko kidea da, baita **BDVA** eta **DAIRO**ko eskubide osoko kidea ere. Claire eta BDVAren bidez, European Partnership on Artificial Intelligence, Data and **Robotics**-en ere parte hartzen du.



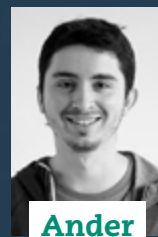
Eneko Agirre

Zuzendaria



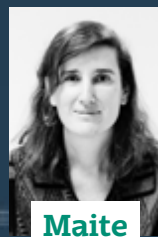
German Rigau

Zuzendariordea



Ander Salaberria

Batzordeko kidea



Maite Oronoz

Batzordeko kidea

Zentroko kideak nazioarteko erreferenteak dira beren jakintza-arloetan. Gaur egun **60 kidek baino gehiagok** osatzen dute zentroa: informatikariek, hizkuntzalariek eta 3 ikerketa-teknikariek. Azken bost urteetan, 200 argitalpen zientifiko baino gehiago argitaratu dituzte zentroa osatzen duten ikertzaileek. Ikaskuntza sakoneko teknikak hizkuntzaren prozesamenduan aplikatzen liderra da taldea, eta azken bi urteetan, **4.000 aipu baino gehiago** izan ditu arlo horretan egin duen lanak. Zentroko kideek aholkulari gisa lagundu dute Espainiako Hizkuntzaren Teknologien Plan Nazionala sortzen, eta, gaur egun, Eusko Jaurlaritza aholkatzen ari dira antzeko lanetan.

CLARIN K CENTRE

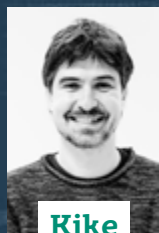


CLAIRE

BDV BIG DATA VALUE ASSOCIATION



ahō LAB



**Kike
Fernandez**

Batzordeko
kidea



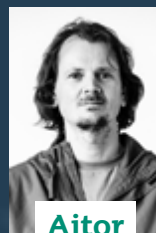
**Itziar
Aldabe**

Batzordeko
kidea



**Inma
Hernaez**

Batzordeko
kidea



**Aitor
Soroa**

Batzordeko
kidea

Bai Ixa bai Aholab errendimendu handiko ikerketa-talde izendatu ditu Eusko Jaurlaritzaren Zientzia Kontseiluak, ikerketa ebaluatzen azken ekitaldian. Historian zehar, taldeek 200 **ikerketa-proiektutan** baino gehiagotan parte hartu dute, eskualde-proiektuetatik hasi eta Europakoetaraino. Halaber, teknologia industriara transferitzeko helburuarekin, 100 **industria-kontratutan** baino gehiagotan parte hartu dute.

Erasmus Mundus+ European Masters Program Hizkuntzaren eta Komunikazioaren Teknologien Europako Masterreko (LCT) kide ere bada HiTZ. Industriak eta ikerketak hizkuntzaren teknologiaren alorrean dituzten

beharrak asetzeko diseinatuta dago Masterra, azkar hazten ari den esparrua baita. HiTZek, halaber, **Doktorego Programa** bat eskaintzen du Hizkuntzaren Análisi eta Prozesamenduz.

Euskal Herriko Unibertsitatea (UPV/EHU) Euskal Autonomia Erkidegoko irakaskuntza- eta ikerketa-erakunde nagusia da. UPV/EHU munduko 400 unibertsitate onenen artean dago, Shanghaiko rankingaren arabera, eta Espainiako Gobernuak Nazioarteko Bikaintasun Campus izendatu du. Euskal Herriko Unibertsitateak, 30 urteko erakunde biziak, 45.000 ikasle ditu, mundu mailako 5.000 langile akademiko eta abangoardiako instalazioak, bere hiru campusetako 20 zentrotan banaturik.

IKERKETA-ARLOAK

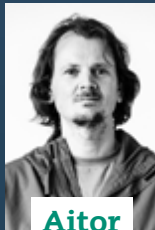
 Ikur honetan klikatuz informazio gehiago eskura dezakezu



Informazioaren Erauzketa eta Berreskurapena



Ikertzaile nagusia:



Aitor Soroa



Itzulpen Automatikoa



Ikertzaile nagusia:



Gorka Labaka



Gizakiaren eta Ordenagailuaren arteko elkarreragina



Ikertzaile nagusia:



Eneko Agirre



Hizketa- eta Hizkuntza-Baliabideak



Ikertzaile nagusia:



Ainara Estarrona

8

HIZKUNTZA TEKNOLOGIA KO EUSKAL ZENTROA

AZPIEGITURA

15

GNU/Linux zerbitzari multi-prozesadore

4

SPARC Solaris zerbitzariak

1

HPC Clusterra, 128 nukleoduna

1

Akustikoki isolaturiko gela, grabazio profesioletarako audio-ekipo batez hornitua



Testuen Analisia



Ikertzaile nagusia:



**Rodrigo
Agerri**



Hizketa- teknologiak



Ikertzaile nagusia:



**Inma
Hernaez**



Hizkuntzaren Prozesamen- dua Medi- kuntzan eta Zuzenbidean



Ikertzaile nagusia:



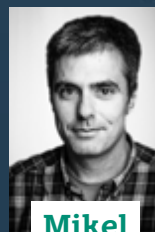
**Arantza
Casillas**



Humanitate Digitalak eta Hezkuntza



Ikertzaile nagusia:



**Mikel
Iruskietza**

hizkuntza

10 zerbitzari,

45

GPUdunak

260

TBtik
gora

sareko biltegiratze
ahalmen gordina

1 1

Behringer 4x4
audio/MIDI
interfazea

Quiet PC
Sentinel
Fanless i10

IKERKUNTZA

ETA TRANSFERENTZIA



34

ikerkuntza
proiektuak

11

ezagutza
transferitzeko
kontratuak

5

bestelako
proiektuak

2

doktoretza-
tesiak
defendaturik

25

aldizkari-
artikuluak
(11 Q1)

36

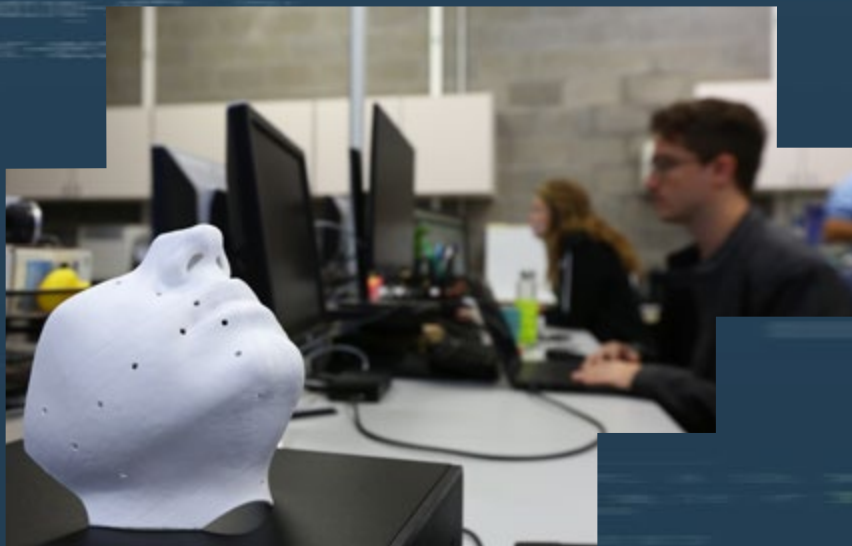
artikuluak
biltzarretan
(8 A or A+)

3

liburu-
kapitulua

10

hitzak



words



36

ikasleak
masterrean

34

ikasleak
doktorego
programan

PRAKTIKAK

9

MALak
amaiturik
HAP/LAPen

75

ikasleak, sakoneko
ikaskuntza bi ikastaro
osagarritan

6

Ikasikerrak

10

barne- eta
kanpo-praktikak

JARDUERAK



40
mintegiak



12

web-mintegiak

7

tailer
antolatuak

7

sariak

Association of Computational Linguistics (ACL) elkarteko Fellow

Informatikako Ikerketa Sari Nazionala

Informatikako Ikertzaile Gaztearen Saria

SEPLN Saria - HNPri buruzko Doktorego Tesi Onenaren Saria

2020 EurAI Doctoral Dissertation Award

SCIE-Fundación BBVA Ikerketa Sariak - Informatika Ikertzaile Gazteak

Ikergazte 2021 - Udalbiltza Sari Berezia

12

HIZKUNTZA TEKNOLOGIAKO EUSKAL ZENTROA

