

CLARIAH-EUS, Gizarte Zientzietako eta Humanitateetako ikertzaile euskaldunentzako sarea

Begoña Altuna, **Jon Alkorta**, Xabier Arregi, Jose Mari Arriola, Ainara Estarrona, Aritz Farwell, Joseba Fernandez de Landa, Xabier Goenaga eta Mikel Iruskieta

Iruña, 2024ko ekainaren 27a

Sarrera eta motibazioa

- Ikertzeko modua etengabe aldatzen ari da
- Teknologia horretan ere eragina
 - Orain arte, teknologia ikerketarako arlo batzuetan bakarrik erabili da
- Gaur egun, Giza eta Gizarte Zientzietan erabiltzen ere hasi da
 - GLAM (galeria, liburutegi, artxibo eta museoek) kultura-ondarearekin
 - Hizkuntzalaritzan, Hizkuntza Teknologia (tresna eta baliabideak) hizkuntza ikertzeko
- Teknologia + Giza eta Gizarte Zientziak: diziplinartekotasuna
- Europako azpiegiturak Humanitate Digitaletan
 - CLARIN: hizkuntza teknologiarik lotuta
 - DARIAH: hezkuntzari, arteari, liburutegiei eta beste alorrei lotuta

CLARIAH-EUSen helburuak

- Giza eta Gizarte Zientzietan euskaraz ikertzeko hizkuntzaren **teknologia garatzea**
 - Tresna eta baliabideen garapena
 - Baliabideen sorkuntza
 - Baliabide horiek CLARIN eta DARIAH azpiegituretan txertatzea
 - Europako azpiegituretan modu egokian garatzeko tresnak sortzea
 - Euskarazko teknologia erabili nahi duten ikertzaileentzako zerbitzua eskaintzea
- Giza eta Gizarte Zientzietan ikertzeko teknologia garatu eta mantenduko duen **ikerketa-komunitatea eratzea**
 - Nazioarteko proiektuetan parte hartzeko ikertzaileen arteko elkarlana sustatu
 - Ikerketa komunitatearen bidez humanitate digitaletan eta hizkuntza teknologian ikuspegi berritzailea sortzea

CLARIAH-EUS

Erakunde babesleak

- Helburua finantziazio-iturri iraunkorrak izatea ikerketa-ekimenak eta baliabideak garatzeko
- Gaur egun, erakunde babesleak
 - Eusko Jaurlaritza (Kultura, eta Hizkuntza Politika saila)
 - Gipuzkoako Foru Aldundia
 - Euskal Herriko Unibertsitatea (UPV/EHU)
 - Euskara, Kultura eta Nazioartekotzearen arloko errektoreordetza
 - HiTZ Hizkuntza Teknologiko Zentroa
 - CLARIAH-EUSeko eta CLARIAH-ESeko bulego teknikoaren arduraduna
- Laguntza horren ondorioz, 4 langile kontratatuta

CLARIAH-EUSen ibilbidea

- **Diseinu fasea (2021-2023)**

- 2021-11-26: CLARIAH-EUSen I. Workshopa

- Helburua 1: arlo desberdineko ikertzaileek dituzten beharrak identifikatzea
 - Helburua 2: azpiegiturak eskain ditzakeen aukera berriak aztertzea
 - Orotara 9 erakundek eta 20 ikertalde desberdinetako 35 ikertzailek parte hartu zuten
 - CLARIAH-EUS azpiegitura babesteko manifestua: 135 ikertzailek eta 9 erakundek sinatuta

- 2021-2023: sarea ehuntzen

- Hainbat erakunde eta ikerketa talderen laguntza bilatzea helburu
 - UPV/EHU, HiTZ zentroa (UPV/EHU), Udako Euskal Unibertsitateko Goi mailako Online Institutua (UEU - GOI), Iker ikerketa-gunea, Elhuyar Fundazioa, EHUko Gogo Elebiduna, Elebilab, Hizkuntzalaritza Teorikorako Taldea (HiTT), Hizkuntzalaritza Diakronikoa, Tipologia eta Euskararen Historia (DLTB), Ixa eta Aholab ikerketa-taldeak, Badalab – Hizkuntza Berrikuntza Laborategia, Soziolinguistika Klusterra eta Giza Eskubideen eta Botere Publikoen Unesco Katedra (UPV/EHU)

CLARIAH- EUSen ibilbidea

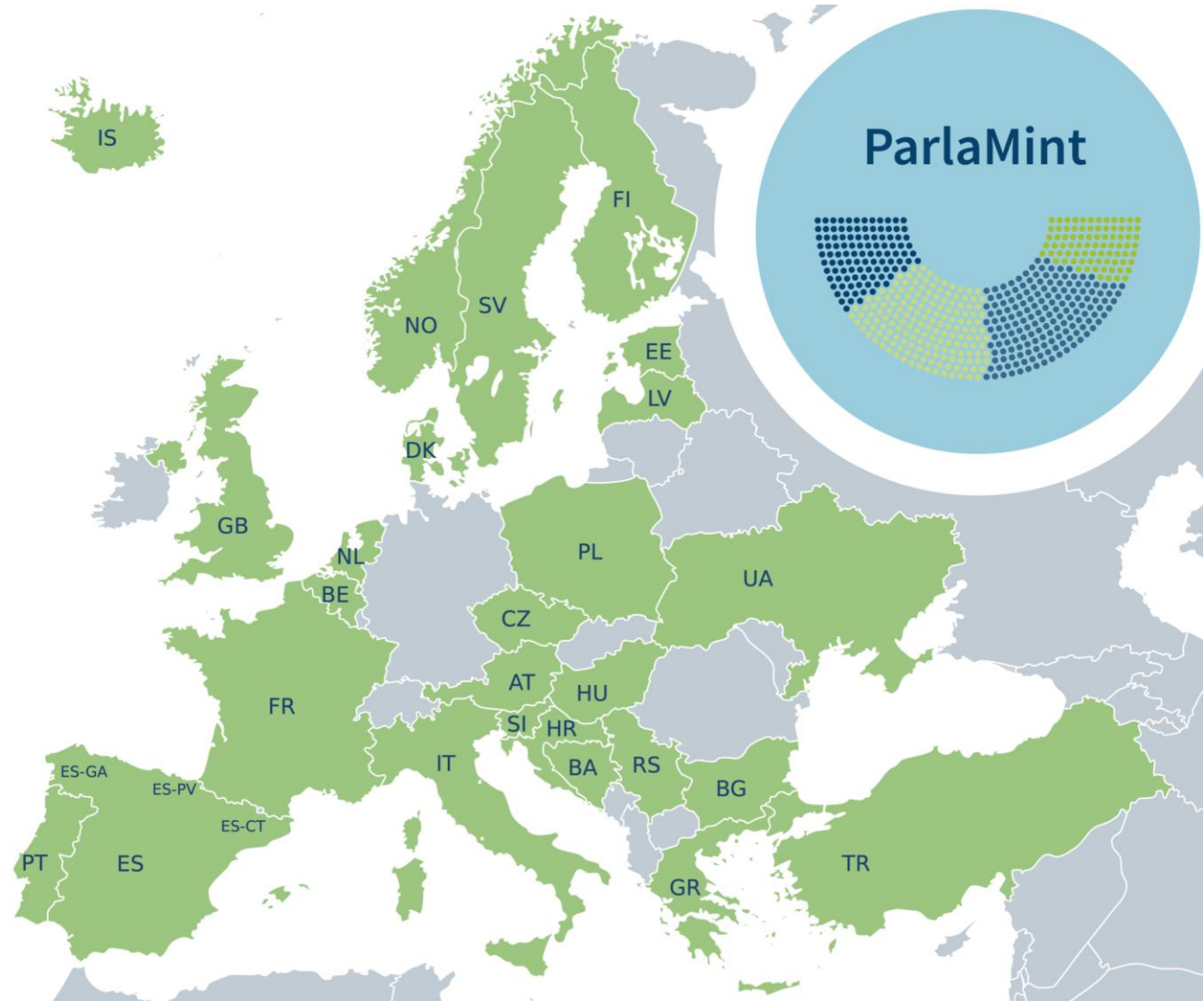
- Akitania Berria-Euskadi-Nafarroa euroeskualdearen baitan: UPNA
- Nevadako unibertsitateko *Center for Basque Studies*
- **Implementazio fasea (2023-2024)**
 - 2023-11-23: II. Workshopa
 - Komunitateko proiektuen berri eman
 - CLARIAH-EUSen lehen batzarra
 - Egitura antolamendua eta bide-orria eztabaidatu
 - 5 urteetarako ildo estrategikoa
 - Eskaera protokoloa garatuta



Kasu praktikoak 1

ParlaMint 4.0

- CLARIAHk sustatuta
- Testuak etiketatuta daude
- Metadatuak ere badaude
- Hainbat adibide:
 - Eusko Legebiltzarra: [https://www.clarin.si/ske/#dashboard?corpname=parlamint40_es_pv].
 - Beste batzuk (Belgika): [https://www.clarin.si/ske/#dashboard?corpname=parlamint21_be].



Kasu praktikoak 2

Gizarte Zientzia Konputazionala

- 3 dataset daude ikertzaileen eskura
 - Helburua: Gizarte Zientzien alorrean esperimentuak egitea eta tresnak garatzea
- **(1) Heldugazte** [<https://github.com/ixa-ehu/heldugazte-corpus>]
- Testu-segida batean agertzen den idazketa-estiloa identifikatzeko
- **(2) Heldugaz-Age corpora** [<https://github.com/joseba-fdl/heldugazte-age-corpus>]
 - Sare sozialetako euskal erabiltzaileen adina asmatzeko dataseta
- **(3) VaxxStance** [<https://vaxxstance.github.io/>]
 - Txertoen inguruan sare sozialetako erabiltzaileek dituzten iritzia identifikatzea

Kasu praktikoak 3

BIM/SAHCOBA

- Morfosintaktikoki etiketatutako euskarazko corpus historikoa
- Helburua: euskal gramatikaren ezaugarri nagusien inguruan azterketa diakroniko sistematikoak egitea
- BIM 1: [<http://ixa2.si.ehu.eus/bim/eu.html>].
- BIM 2: [<http://bim.ixaeus/search#>]



*Sintaktikoki etiketatutako
euskarazko corpus historikoa*



SAHCOBA

*Syntactically Annotated Historical Corpus in Basque
Morfosintaktikoki etiketatutako euskarazko corpus historikoa*

CLARIAH-EUSen etorkizuna

- Izaera instituzionala

- CLARIAH-EUS CLARIN B-zentroa bihurtzea [<https://www.clarin.eu/content/certified-b-centres>]
 - Ikertzaileei zerbitzu teknikoa eta orientazio didaktikoa eskaintzeko

- Lan-ildo nagusiak

- CLARIAH-EUSeko ikertzaileentzat baliabideak eta zerbitzuak sortu eta egokitu
- Euskal komunitatean estrategikoki beharrezkoak diren baliabideak eta zerbitzuak sortu edo egokitu
- CLARINen eta DARIAHan txerta daitezkeen baliabideak edo zerbitzuak sortu edo egokitu