



HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANA

2021-2024



EUSKO JAURLARITZA
GOBIERNO VASCO

KULTURA ETA HIZKUNTZA
POLITIKA SAILA
DEPARTAMENTO DE CULTURA
Y POLÍTICA LINGÜÍSTICA



AURKIBIDEA

Hitzaurrea	3
1. Sarrera. Motibazioa	5
1.1 Hizkuntza-teknologiak eta Administrazio publikoa	5
1.2 Euskara Hizkuntza-teknologietan integratzea	6
1.3 Zergatik Hizkuntza-teknologiaren Ekintza-plana?	7
2. Bisioa	10
3. Helburu nagusiak	11
4. Ekintza-plana	13
4.1 Ekintza-planaren ardatzak	13
4.1.1 Kudeaketa eta printzipioak	13
4.1.2. Hizkuntza-baliabideak: testu zein ahozko materialak biltzea, tratatzea eta etiketatzea	14
4.1.3. Hizkuntza prozesatzeko lanabesak	15
4.1.4. Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea	15
4.1.5. Itzulpen automatikoa	16
4.1.6. Hizketaren tratamendua: ahots-interfazeak	16
4.2 Ekintza-planaren arlo estrategikoak	17
4.2.1 Administrazio elektronikoa	18
4.2.2 Hezkuntza	19
4.2.3 Helduen euskalduntzea	20
4.2.4 Osasungintza	21
4.2.5 Justizia	22
4.2.6 Turismoa	23
4.2.7 Kultura, hedabideak, humanitate digitalak, liburutegi digitalak...	24
4.2.8 Segurtasuna	25
4.3 Euskarazko Hizkuntza-teknologiaren Ekintza-plana, 2021-2024: itsasargi-proiektuak	26
5. Ekintza-planaren hasierako aurrekontu-aurreikuspena	52
6. Hizkuntza-teknologiaren ekintza-planaren jarraipena	53
7. Erreferentziak	57



HITZAURREA

Azken urteotan ikusi dugu zein garrantzitsua den euskararentzat Adimen artifizialaren garapenean integratuta egotea. ITZULI itzultzaile neuronalak da adibide esanguratsuena, horren bidez argi eta garbi erakutsi baita euskarazko kalitateko zerbitzuak eskaini dakizkiekeela bai administrazio publikoei bai gizarte osoari.

Hain zuzen, Hizkuntza-teknologietan aurrera egiteko eta euskara garapen teknologiko handiko Europako hizkuntzen artean jartzeko, **GAITU 2021-2024 Hizkuntza-teknologien Ekintza-plana** egin dugu. Ekintza-plana da Administrazioaren proiektu zehatzez osatu delako. Izan ere, gure ustez, ezinbestekoa da, batetik, Eusko Jaurlaritzan eta bere erakunde autonomiadunetan, legegintzaldi honetan jada abiatuta dauden proiektuak eta abiatuko direnak jasotzea eta horien berri ematea, eta, bestetik, euskara Hizkuntza-teknologietan modu eraginkorrean integratzeko, aldi berean garatzea hizkuntza-baliabideak (datuak eta hizkuntza-ereduak), hizkuntza tratatzeko lanabesak (analizatzaileak, etiketatzaileak, anonimizatzaileak) eta tresnak, eta horien erabilera sustatzea. Horrela, Administrazioaren benetako premiei erantzuteaz gain, aurrerantzean, Hizkuntza-teknologietan euskara integratzen jarraitzeko beharrezkoa den azpiegitura linguistikoa eratzen joango gara.

Ekintza-plan hau lehenengo urratsa da, abiapuntua. Aipatu moduan, oraingoan, Administrazioaren proiektuen argazkia erakutsi nahi dugu. Ez da, beraz, jaso Administrazioetik kanpoko proiekturik, ikerketakorik nahiz bestelakorik. Horiek hurrengo fase batean jasoko dira. Dena den, lehenengo urrats honetan, abian dauden proiektuetan, Administrazioarekin batera, zenbait aktore daude inplikaturik, jada, eta egongo dira datozen urteotan ere: eduki-sortzaileak, ikerketa-taldeak, enpresa garatzaileak eta enpresa erabiltzaileak. Izan ere, ekimen publikotik apustu segurua egin nahi da euskarazko Hizkuntza-teknologien alde eta horiek garatzeko behar diren hizkuntza-datu masiboak batzearen alde. Hori dela eta, ekintza-plan honetan dagoen proiektuetako bat da 'Datuaren Bulegoa' sortzea, hain zuzen, jada erabilgarri dauden datuak eta eredu linguistikoak, eta etorkizunean egongo direnak, bertan biltzeko eta sektoreko enpresentzat, ikerlariarentzat eta gainerako agenteentzat eskuragarri jartzeko.

Administrazioak trakzio-lana egin behar du eta, hortaz, ekimen honi esker, euskarazko kalitateko Hizkuntza-teknologien bidez, oinarritzko hizkuntza-zerbitzu transbertsalak (itzulpen automatiko neuronalak, hizketa-ereduak, hizkuntza-ereduak eta datuak) garatu eta eskaini nahi

ditu, lehenengo eta behin administrazio publiko guztietan erabiltzeko eta, azken finean, herritarrei zerbitzu publiko hobek eskaintzeko.

Ekintza-plana irekia da eta bertan jasotako proiektuen jarraipena egingo da. Gainera, legegintzaldian zehar, proiektu berriak gehituko dira eta, legegintzaldiaren amaieran, ebaluazioa egingo da, hurrengo faseari ekin aurretik. Horrela, informazio zehatza izango dugu: erabilitako teknologia, garatutako datu, lanabes eta tresna linguistikoak eta haien berrerabilpena, erabilitako aurrekontua eta garatzen jarraitzeko behar-ko litzatekeena.

Esan behar da Hizkuntza-teknologiaren Ekintza-plan hau Ekonomia-aren Garapen, Jasangarritasun eta Ingurumen Sailak, 2021ean onartu zuen “Euskadiko Eraldaketa Digitalerako Estrategia 2025” delakoarekin uztartzen dela eta horren garapena ere badela, Adimen artifiziala baita bertan aipatzen den palanka teknologikoetako bat eta, gainera, planean jasotako proiektuek aipatu estrategian jasotzen diren aplikazio-esparruen premiei erantzuten baitiete.

Halaber, ekintza-plan honek Gobernantza Publiko eta Autogobernu Sailak 2021ean argitaratutako “ARDATZ Gobernantzaren, Berrikuntza Publikoaren eta Gobernu Digitalaren Plan Estrategikoa 2030” delakoarekin bat egiten du hainbat ekimen partekatuak dituztelako bi planek. Eta ezin aipatu gabe utzi Eusko Jaurlaritzaren Lehendakariaren “Zientzia, Teknologia eta Berrikuntzako Plana Euskadi 2030”; azken finean bi plan horiek eta ekintza-plan hau elkarren osagarriak dira eta helburu komunak dituzte.

Azkenik, Ekintza-planean agertzen diren proiektuak direla eta, Estatu-ko Gobernuak 2015ean argitaratu zuen “*Plan de Impulso de las Tecnologías del Lenguaje*” planarekiko lotura egiten saiatu gara.

Nire esker ona Hizkuntza-teknologiaren Ekintza-plana egiteko informazioa eman diguten guztioi!

Vitoria-Gasteiz, 2022ko otsailaren 25a

Bingen Zupiria Gorostidi
Kultura eta Hizkuntza Politika Sailburua

1. SARRERA. MOTIBAZIOA

1.1 HIZKUNTZA-TEKNOLOGIAK ETA ADMINISTRAZIO PUBLIKOA

Gaurko mundua teknologiei esker interkonektatuta dago eta, horretan, Hizkuntza-teknologiek izugarriko garrantzia hartu dute. Premia eta behar berriak sortu dira, beti gizarteari eta ekonomiari lotuak eta, hori horrela, teknologia, hizkuntzei lotuta eta, azken urteotan, Adimen artifizialari esker, ezinbesteko faktore bihurtu da gizakion bizitzetan.

Hizkuntza-teknologiaren helburua hizkuntza prozesatzea da, idatzia nahiz ahozkoa. Adimen artifizialaren atalak diren Hizkuntzaren Prozesamendua (HP) eta Itzulpen Automatikoa (IA) dira Hizkuntza-teknologiak. Testuetan jasotako datu masiboak ustiatzen dituzte, baita webguneetakoak eta sare sozialetakoak ere. Prozesadore linguistikoen bitartez, besteak beste, aplikazio hauek eskaintzen dira gaur egun: izendatutako entitateak hautematea (izen propioak, datak, markak, produktuak, eta abar), dokumentuen bilaketa adimenduna, dokumentuak sailkatzea eta multzokatzea, laburpen automatikoak egitea, dokumentuetatik datuak erazte (sentimendu-analisia eta iritzi-meatarritza), sare sozialetako erreputazioaren jarraipena eta monitorizazioa, alertak sortzea, testu-zuzentzaileak, hizkuntzak ikasteko sistemak, galderak automatikoki erantzuteko sistemak, eta testuak modu automatikoan itzultzea.

Teknologia horien merkatua azkar hazten ari da eta hurrengo urteetarako hazkuntza handiagoa aurreikusten da.

Administrazio publikoek hizkuntza naturala prozesatzeko teknologiak, baita itzulpen automatikoarenak ere, haien eguneroko jardunean txertatu behar dituzte, zerbitzu publikoaren kalitatea eta gaitasuna hobetzeko. Zenbait arlotan, gainera, hala nola osasunean edo justizian, kudeatzen dituzten edukiak direla eta, hizkuntza-teknologiak erabakigarriak dira eta, hortaz, ulermen sakonagoan eta itzulpen automatikoan oinarritutako zerbitzu berriak garatu behar dira arlo horietarako.

Bestalde, hizkuntza-teknologiek administrazioarekiko eskaria handitu eta trakzio gisa jardun dezakete eta, horrela, sektorearen garapena azkartu. Administrazioaren inplikazioak ikuspenera, ospea eta bermea ematen dizkiete garatutako aplikazioei eta produktuei etekina ateratzeko. Halaber, administrazioen parte-hartzeak aukera ematen du euskararen baliabideak eta prozesadoreak irekiak eta partekatuak izateko, eta datu publikoak berrerabiltzean oinarritutako berrikuntza sustatzeko.

Administrazioen arteko interakzioa ere areagotu egiten da teknologiei esker. Gure testuinguruari dagokionez, hiru administrazio (orokorra, forala eta tokikoa) ditugu, eta horietan guztietan, hizkuntza-teknologiak oso lagungarriak dira.

Azken finean, administrazio publikoek hizkuntza-teknologiaren bultzatzaile izan behar dute; gainera, plataforma bateratuak sortu behar dituzte baliabideak eta produktuak bertatik eskaintzeko eta partekatzeko erabiltzaile desberdinekin.

1.2 EUSKARA HIZKUNTZA-TEKNOLOGIETAN INTEGRATZEA

Indartsu datoz Adimen artifiziala eta hizkuntza-teknologiak eta abagune honek politika ondo diseinatuak eta eraginkorrak behar ditu. Estatu ahaltsuek eta, batez ere, erraldoi teknologikoek —Google, Microsoft, Facebook, Amazon eta IBM bezalakoek— lekua hartu dute teknologia horiei dagokienez, uhin horrek ihes egin ez diezaien. Jokaleku globalean gertatzen ari da eraldaketa teknologikoa, baina ez ordea modu homogeneoan.

Hizkuntza-teknologiak aktibatzen dituzten politikak ari garelako, kontuan hartu beharreko ezaugarria da heterogeneotasuna. Batetik, arloko tresnen eta aplikazioen errendimendua eta erabilgarritasuna ez da inola ere uniforme, denak hizkuntza-teknologiaren barnekoak izanagatik ere. Bestetik, alde handia dago hizkuntza batzuen eta besteen garapenean, edo, zehatzago esanda, hizkuntza boteretsuen eta eremu urriko hizkuntzen artean.

Badirudi errealitate anizkoitz eta konplexu horrek alderdi teknologikoa ez ezik, alderdi ekonomikoa eta soziala ere astinduko duela. Hortaz, ezinbestekoa da erakundearen, bereziki erakunde publikoen, esku-hartze ordenatua eta eraginkorra.

Euskararen esparrura etorritik, arrisku handia dago hizkuntza hegemonikoek guztiz berengana dezaten arreta eta sostengua, eta hizkuntza minorizatuak, euskara barne, zokoratuta gera daitezkeen.

Hizkuntza-teknologiaren barruan, azken urteotan gertatu diren hobekuntza handiek (*big data* delakoarekin eta sare neuronalekin, besteak beste) bigarren susperraldi bat eragin dezakete euskararen berreskurapen-prozesuan. Kualitatiboki askoz handiagoa izan daiteke bigarren bultzada hori gure gizartean. Hizketaren tratamenduan, itzulpen automatikoan edo testuen analisisetan gertatu diren hobekuntza nabarmenek modu esanguratsuan lagundu dezakete euskararen erabilera errazten eta zabaltzen.

Euskara hizkuntza-teknologiaren eta ingurune digitalean integratzea beti izan da helburu estrategikoa Eusko Jaurlaritzarentzat. Hori dela eta, Gobernu-programetan, agenda estrategikoetan, landu eta onartutako plan guztietan, lekua izan dute teknologia horiek. Modu askotara sustatu izan ditu Jaurlaritzak hizkuntza-teknologiak azken urteotan: dirulaguntzak, ikerketa-programak eta produktu jakin batzuk garatuz erabiltzaileen eskura jartzeko, adibidez, itzultzaile automatikoa, lehenago itzultzaile gramatikala eta, orain bi urte, Adimen artifizialari esker, ITZULI itzultzaile neuronal (www.euskadi.eus/itzuli), ezinbesteko tresna bihurtu dena. Edo zenbait hizkuntza-baliabide eta tresna biltzen dituen ITZULI+ tresna-barra (www.euskadi.eus/itzuliplus), irakaskuntzan nahiz edozein lanetan oso erabilgarria.

Baina, denbora galdu gabe, tresna berriak garatzen jarraitu behar da bermatzeko euskal komunitateak egun merkatuan dauden eta etorkizunean egongo diren produktuak euskaraz kontsumitu ahal izango dituela: ahotsaren sintesia eta ezagutza, automatikoki azpigitulatuak ikus-entzu-

nezko edukiak, ahots-laguntzaileak, zerbitzu digitalen plataformak, bideojokoak, mediku-zerbitzuak, itzultzaile neuronal berezituak, ahots-ahots itzulpen automatikoa, galdera-erantzuna el-karrizketa-sistemak...).

1.3 ZERGATIK HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANA?

Euskara ez da garapen teknolinguistikoa beharrari aurre egin behar dion hizkuntza bakarra. META-NETek¹, 2013an egindako azterketan, aurreikusi zuen gutxienez Europako 21 hizkuntza digitalki desagertzeko arriskuan zeudela. Nahiz eta META-NETek euskara hizkuntza horietako bat zela esan, aipatu behar da badaudela hizkuntza handiagoak, askoz hiztun-kopuru handiagoa dutenak, eta horiek ere erronka berdintsuei aurre egin behar dietela. Izan ere, gaur egun ingelesa jaun eta jabe den munduan euskarak ingurune digitalaren euskarria ezinbestekoa du.

Europar Batasunean dagoen hizkuntza-kopuru handia egotearen ondorioz sortzen diren hesiak gainditzeko aukera paregabea ikusi du Europako Batzordeak hizkuntza-teknologietan. Hori dela eta, azken urteotan gauzak asko mugitzen ari dira Europan, nahiz eta bertan ofizialak ez diren hizkuntzak ez diren zuzenean kontuan hartzen. Dena den, erreferentzia eta aukera inportanteak izan daitezke euskararentzat ere han landu diren estrategiak eta dokumentuak: *Digital Single Market*² (Merkatu Digital Bakarra), *Language equality in the digital age*³ (Hizkuntza-berdintasuna aro digitalean) edota *Artificial Intelligence: A European Perspective*⁴ (Adimen artifiziala: europar ikuspegia). 2018an Europako Parlamentuak hauxe esan zuen:

“The European Parliament calls on the Commission and the Member States to develop strategies and policy action to facilitate multilingualism in the digital market; requests, in this context, that the Commission and the Member States define the minimum language resources that all European languages should possess, such as data sets, lexicons, speech records, translation memories, annotated corpora and encyclopaedic content, in order to prevent digital extinction”.

(Europako Parlamentuak dei egiten die Europako Batzordeari eta estatukideei estrategiak eta ekintza politikoak gara ditzaten, merkatu digitalean hizkuntza-aniztasuna errazteko; eskatzen die, ildo horretatik, Batzordeari eta estatu kideei zehaztu ditzatela Europako hizkuntza guztiek eduki behar dituzten gutxienerako hizkuntza-baliabideak, hala nola, datuak, lexikoak, ahozko grabazioak, itzulpen-memoriak, corpus etiketatuak eta eduki entziklopedikoak, iraungipen digitala eragozteko).

Bestalde, Espainiako Gobernuak 2015ean *Plan de Impulso de las Tecnologías del Lenguaje*⁵ izeneko plana onartu zuen. Honela zehazten da planaren helburua:

1 <https://link.springer.com/content/pdf/bfm%3A978-3-642-30796-6%2F1.pdf>

2 <https://ec.europa.eu/digital-single-market/en/shaping-digital-single-market>

3 http://www.europarl.europa.eu/thinktank/es/document.html?reference=EPRS_STU%282017%29598621

4 <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/artificial-intelligence-european-perspective>

5 <https://plantl.mineco.gob.es/tecnologias-lenguaje/PTL/Bibliotecaimpulsotecnologiaslenguaje/Detalle%20del%20Plan/Plan-Impulso-Tecnologias-Lenguaje.pdf>

El Plan de Impulso de las Tecnologías del Lenguaje (Plan TL) tiene como objetivo fomentar el desarrollo del procesamiento del lenguaje natural, la traducción automática y los sistemas conversacionales en lengua española y lenguas cooficiales.

(Hizkuntza-teknologiak sustatzeko Planak (TL Plana) hizkuntza naturalaren prozesamendua, itzulpen automatikoa eta elkarrizketa-sistemen garapena sustatzea du helburu, gaztelaniaz eta hizkuntza koofizialetan).

Planak hauxe ere ezartzen du:

Es un Plan interdepartamental dado el carácter multidisciplinar de las tecnologías del lenguaje, y tiene como premisa coordinar todas las actuaciones de la Administración General del Estado, en coordinación con las Comunidades Autónomas, para el impulso de las tecnologías del lenguaje.

(Sailen arteko plana da, hizkuntza-teknologiek diziplina anitzeko izaera baitute, eta Estatuko Administrazio Orokorren jarduera guztiak koordinatuko ditu, autonomia-erkidegoekin batera, hizkuntza-teknologiak sustatzeko).

Bada, Hizkuntza-teknologiaren ekintza-plan hau Euskadiko Administrazio Orokorreko sail eta erakunde autonomiadunen artean garatzekoa izan behar da eta Espainiako Gobernuaren planarekin koordinatzea komeniko litzateke, ahaleginak eta sinergiak uztartzeko eta bikoiztasunak ekiditeko.

Eusko Jaurlaritzak, bestalde, beti izan du oso presente euskararentzat ezinbestekoa dela Ingurune digitalean eta Hizkuntza-teknologietan integratzea. Kultura eta Hizkuntza Politika Sailak kudeatzen duen *Euskararen Aholku Batzordearen Ingurune Digitalean Euskara Sustatzeko Batzorde-atal Bereziak landutako* “Euskararen Ingurune Digitala, Gomendioak 2021-2024” argitalpenaren 4.1 ekintza hauxe da: **“Hizkuntza-teknologiaren plana zehaztu eta idatzi”**.

Batzorde-atal berezi horrek ere landu eta 2016an Kultura eta Hizkuntza Politika Sailak argitaratu zuen “Gomendioak herri-aginteentzat” lanean, jada, eman zen gomendio hori: “Hizkuntza-teknologiak garatzen lagundu eta berariazko plana garatu”. Hain zuzen, gomendio horri ahalik eta lasterren erantzun nahian eta Hizkuntza-teknologiaren plan bat egite aldera, Kultura eta Hizkuntza Politika Sailak Euskal Herriko Unibertsitateko Hitz Zentroari txosten bat enkargatu zion 2018an. Horren emaitza da 2020ko otsaileko “Hizkuntza-teknologiak sustatzeko planaren txosten tekniko”. Bertan jasotakoak oso kontuan izan ditugu ekintza-plan hau egiteko orduan.

Txostenean zehazten denaren arabera, kalitate handiko hizkuntza-teknologia garatzea ezinbestekoa eta presazkoa da euskararen kasuan. Ezinbestekoa, hizkuntzaren beraren irautea bermatu nahi bada; eta presazkoa, inondik ere, teknologia horien bilakabidea eta berritze-abiadura ikusita atzean geratzeko arriskua benetakoa baita.

Bestalde, Hizkuntza-teknologiaren Ekintza-plana Eusko Jaurlaritzak 2021ean onartu eta argitaratu dituen bi plan garrantzitsurekin uztartu behar da:

- Ekonomiaren Garapen, Jasangarritasun eta Ingurumen Sailaren “Euskadiko Eraldaketa Digitalerako Estrategia 2025”. Ekintza-plan hau estrategia horren barruan kokatu behar da; izan ere, bertan, Adimen artifiziala da palanka teknologietako bat eta, zehaztutako aplikazio-esparruen artean, ekintza-plan honetan estrategikotzat jotako eremuak agertzen dira, besteak beste: e-Administrazioa, e-Osasuna eta ongizatea, e-Justizia, Segurtasuna,

Hezkuntza, Turismoa. Beraz, estrategia orokor horren garapen moduan ulertu behar da Ekintza-plana.

- Gobernantza Publiko eta Autogobernu Sailaren “ARDATZ Gobernantzaren, Berrikuntza Publikoaren eta Gobernu Digitalaren Plan Estrategikoa 2030”. Plan horren 4.3 atalean jaso dira estrategiak eta ardatzak: “I. Gobernu zuzena, berdinzalea eta elebiduna (euskara, lan eta zerbitzu hizkuntza); II. Gobernu berritzailea eta etengabeko ikaskuntzan diharduena; III. Herritarrak ardatz dituen gobernu digitala; IV. Datuetan oinarritutako Gobernuak ematen dituen (Datuen eta informazioaren gobernantza); V. Gobernu irekia. Aipatu bezala, Hizkuntza-teknologiaren Ekintza-plana sailarteko plana da eta, hain zuzen, ARDATZ planean jasotako helburu batzuk erdiesten lagunduko du, bertan jasotako proiektu batzuk Gobernantza Publiko eta Autogobernu Sailarenak baitira.

Azkenik, ekintza-plan hau hasierako urratsa da (2021-2024 epealdirako izango da) euskal administrazio orokorraren ohiko jardueran euskarazko hizkuntza-teknologiak modu transbertsalean integratzeari begira. Ekintza-plana izango da, Eusko Jaurlaritzan Hizkuntza-teknologiak erabiltzea, garatzen ari diren eta legegintzaldian zehar garatuko diren proiektuak jasoko dituelako. Teknologiak azkar doaz eta, garrantzitsua da ahalik eta lasterren, erabiltzaileei produktu zehatzak eskaintzea eta, aldi berean, produktu horiek garatzeko erabiliko diren hizkuntza-corporusak eta baliabideak berrerabiltzeko gordetzea eta plataforma batean jartzea.

Ekintza-plan honen atzetik beste batzuk etorriko dira. Oraingo honen abiapuntua hauxe da: Eusko Jaurlaritzan, jada, martxan dauden proiektuak eta 2021-2024 legegintzaldian zehar egongo direnak jasotzea eta horien jarraipena egitea.

Hizkuntza-teknologiak dira Ekintza-planaren ardatza; beraz, ez dira jaso bertan ingurune digitaleko ekimenak, horiek Euskararen Aholku Batzordearen *Ingurune Digitalean Euskara Sustatzeko Batzorde-atal Bereziaren* esku geratu baitira. Ez dira jaso, halaber, ikerketa-proiektuak, horiek aurrerago txertatuko baitira.

Ikerketa dela eta, esan behar da 2021ean Eusko Jaurlaritzako Lehendakariak 2021ean argitaratutako “Zientzia, Teknologia eta Berrikuntzako Plana Euskadi 2030” plana zuzenean lotzen dela Euskal Autonomia Erkidegoak ikerketaren eta berrikuntzaren alde egiten duen apustu estrategikoarekin.

2. BISIOA

Euskal Autonomia Erkidegoko Administrazio Orokorren barruan zerbitzu publikoaren kalitatea eta gaitasuna hobetzea, Hizkuntza-teknologiak (Adimen artifizialaren bidez, hizkuntza naturala prozesatzeko teknologiak eta itzulpen automatikoa) txertatuz, eta, gainera, eskariaren bultzatzaile edo trakzio gisa jardunez. Administrazioaren kudeaketa publiko elebidunaren jardueraren testuinguruan sortutako hizkuntza-baliabideak biltzea, estandarizatzea, zabaltzea eta berrerabiltea sustatu nahi da ekintza-plan honen bidez.

Etorkizunari begira Hizkuntza-teknologiak oso garrantzitsuak izango dira euskararentzat. Gure txikitasunaren esparruan Administrazio publikoak bidea erakutsi behar du teknologia horiek erabat integratzen eta bultzatzen, neurri handi batean herritarrekiko interakzioa bermatzeko eta, ondorioz, ziurtatzeko euskara ahalik eta goera gehienetan erabiliko dela.

Ekintza-plan honen bidez Hizkuntza-teknologiei lehentasuna eman nahi zaie hori baita hizkuntzaren biziraupen teknologikoa bermatzeko modu bakarra. Teknologia horien garapena gerezina da eta, aukera baliatzen ez badugu, beste hizkuntza batzuek (ingeleza, gaztelania) hartuko dute espazio hori.

Hizkuntza-teknologiak berrikuntzari lotutako gorantz doan zeharkako sektorea dira eta hazkundea, lehiakortasuna eta kalitateko enplegua sustatzeko gaitasuna dute.

Beharrezkoa da Administrazio Publikoaren barruan ekintzak bultzatzea eta koordinatzea, sinergiak bilatzea eta ahaleginak bikoiztea eragozte. Hizkuntza-teknologiek (Adimen artifiziala) azpiegitura informatiko ahaltsua (superkonputazioa) behar dute eta, printzipioz, Administrazio Publikoak horretan inbertsioak egiteko aukera handiagoa izaten du.

Ekintza-plan hau luzaroan jarraipena izango duen plan luze baten hasiera da. Administrazio Publikoak bidea jorratu behar du etorkizuneko belaunaldientzat azpiegitura iraunkorra bermatzeko, euskara, benetan, garapen teknologikoa dela eta, berrikuntzaren erdigunean egon dadin.

Azken finean, Administrazioaren helburua pertsonen bizimodua hobetzea izan behar da eta, horretarako, gaur egun, nahitaez, mota guztietako teknologiak ezinbestekoak dira.

3. HELBURU NAGUSIAK

Hauek dira Hizkuntza-teknologiaren Ekintza-planaren helburuak:

- Euskadiko Administrazio Orokorrean oinarriak jartzea Hizkuntza-teknologiak arlo estrategikoetan (Administrazio elektronikoa, Hezkuntza, Helduen Euskalduntzea, Osasungintza, Justizia, Turismoa, Kultura eta Hedabideak, eta Segurtasuna), modu transbertsalean, garatzen eta txertatzen joateko, horrela, Adimen artifizialaren ekosistema sortzeko baldintzak jarritz.
- Administrazio Orokorrearen barruan Hizkuntza-teknologiaren garapen eta aplikazioari dagokienez, lankidetzaren eta koordinazioaren bultzatzea, sinergiak aprobetxatzeko, ekintzak sinkronizatzeke, erredundantziak saihesteko eta, azken finean, herritarrei zerbitzu publiko hobea emateko.
- Itsasargi-proiektuak edo proiektu transbertsalak bideratzea eta gauzatzea administrazio publikoak traxzioa izan behar baitira.
- Euskarak kalitateko hizkuntza-teknologiak garatzeko behar dituen baliabideak sortu eta agente interesdun guztien eskura jartzeko bidea egitea: testu zein ahozko corpusak (gordinak, etiketatutak, orokorrak eta espezializatuak, elebakarrak, elebidunak eta eleaniztunak). Horretarako, datuen plataforma eta datuen erabilerarako protokoloa sortuko dira.
- Gure administrazio publikoa elebiduna izanda, garapen teknolinguistikoari dagokionez, gaztelania eta euskararen artean dagoen distantzia murriztea.
- Hizkuntza-teknologiak erabilia, EAEko Administrazio Orokorreko arlo estrategikoetan martxan dauden proiektuak eta 2021-2024 aldiaren garatuko direnak jasotzea eta ezagutaraztea, baita arlo estrategiko horietan edo beste batzuetan etorkizunean garatu beharko liratekeen proposamenak ere.
- Administrazioen eta, oro har, herritarren aurrean Hizkuntza-teknologiak ikusgai jartzea, ikus dezaten teknologia horiek haien ongizatean duten eragin positiboa.
- Euskal Autonomia Erkidegoko Hizkuntza-teknologiaren industriaren sektorea indartzea Administrazioarekiko lankidetzaren bidez, eta, hortaz, lankidetzaren publiko-pribatua bultzatuz eta indartuz.
- Euskarari dagokionez, Adimen artifizialaren garapen jasangarria, integratzailea, segurua eta bizidun sustatzea. Gainera, gaur egun hizkuntza-teknologiek duten booma aprobetxatu behar da euskara hori egon dadin eta, horrela, beste hizkuntzetan garatuta dauden tresnak, adibidez, elkarrizketa-laguntzaileak eta chatbot-ak, euskaldunek ere eskura eta erabilgarri izango dituzte.



- Ekintza-planaren koordinaziorako eta jarraipenerako mekanismoak ezartzea. Plana dinamikoa (irekia) eta moldagarria izango da eta egoeraren bilakaerara egokitu beharko da, izandako esperientziez baliatuta, aurrera begira. Halaber, lagunduko du arlo desberdinetan dauden premiak eta beharrak atzematen.

4. EKINTZA-PLANA

4.1 EKINTZA-PLANAREN ARDATZAK

4.1.1 Kudeaketa eta printzipioak

Euskadiko Administrazio Orokorraren arlo estrategikoetan Hizkuntza-teknologiak modu transbertsean, garatzen eta txertatzen joateko oinarriak jartzea denez Ekintza-planaren helburu nagusietako bat, Eusko Jaurlaritzak kudeatuko du. Sailen arteko plana izanda, Hizkuntza Politikarako Sailburuordetzak koordinatuko du, Ekintza-planean estrategikotzat jotako eremuetako arduradunen laguntzaz. Elkarlana bultzatzea sustatuko da, estrategia komunak bilatuz, ideiak trukatzuz, baliabideak eta tresnak partekatuz, premiak eta beharrak elkarrekin aztertuz eta lane-rako bideak proposatuz.

Printzipioak:

1. Parte-hartzea: ekintza-plan honetan inplikaturako Administrazio arlo estrategikoez gain, gainerako sailen partaidetza bultzatuko da. Halaber, gainerako administrazio publikoena ere bilatuko da.
2. Gardentasuna: 2021-2024 legegintzaldian, Administrazioan garatzen ari diren eta garatuko diren proiektuak eta horien jarraipenaren gaineko informazioa ikusgai egongo dira.
3. Irekitasuna eta zabalkundea: Administrazioak sustatutako proiektuetatik sortutako datuak, baliabideak eta tresnak erabiltzaileen eta interesdunen eskura jarriko dira. Horiek berrerabil-tzea izango da Ekintza-planaren jomugetakako bat.
4. Iraunkortasuna: ekintza-plan hau abiapuntua da. 2021-2024 legegintzaldian garatzen ari di-ren eta garatuko diren proiektuak, euskararen hizkuntza-teknologiak integratzen dituztenak, jasotzen ditu. Lehenengo aldiz, proiektu horiek guztiak dokumentu batean jaso dira. Proiektu horiek eta aurrerantzean etorriko diren berriek jarraipena behar dute izan. Beraz, legegintzal-diz legegintzaldi Ekintza-plana garatzen jarraitu beharko da.
5. Dinamikotasuna eta moldagarritasuna: Ekintza-plana ez da dokumentu itxia, hain aldakorrek diren teknologiek eraginda, moldaketak izan ditzake (proiektuetan aldaketak sartzeko edota proiektu berriak integratzeko).

6. Zeharkakotasuna: Ekintza-planak, hasiera batean, Eusko Jaurlaritzako sail batzuk (estrategikotzat jotakoak) inplikatu baditu ere, gainerakoetara zabalduko da. Halaber, Ekonomiaren Garapen, Jasangarritasun eta Ingurumen Sailaren “Euskadiko Eraldaketa Digitalerako Estrategia 2025” eta Gobernantza Publiko eta Autogobernu Sailaren “ARDATZ Gobernantzaren, Berrikuntza Publikoaren eta Gobernu Digitalaren Plan Estrategikoa 2030” planekin uztartuko da. Gainera, “*Plan de Impulso de las Tecnologías del Lenguaje*” delakoarekin ere lotuko da.
7. Erakunde publikoen arteko lankidetzak: haien digitalizazioa bultzatzeko, sinergiak aprobetxatzeko, baliabideak eta tresnak partekatzeko eta, azken finean, Administrazioaren maila guztietan, herritarrei euskaraz kalitateko zerbitzu publikoak eskaintzeko. Izan ere, zerbitzu publikoak ez dira onak izango euskaraz ere eskaintzen ez badira.
8. Lankidetzak publiko-privatuak: proiektuak elkarrekin garatuz, sektorea indartuko da, teknologiei nahiz ekonomiari begiratuta.
9. Hizkuntza-teknologiaren bidez onura orokorra lortzea gizarte osoarentzat.

4.1.2. Hizkuntza-baliabideak: testu zein ahozko materialak biltzea, tratatzea eta etiketatzea

Hizkuntza-teknologiak garatzeko ezinbestekoak diren kalitateko baliabide linguistikoak bildu, tratatu, etiketatu eta plataforma publiko batean jaso behar dira, formatu egokietan eta lizentzia librepean. Horrekin batera, baliabide horien erabilera desberdinak aintzakotzat hartuta, protokolo bat definitu eta zehaztu behar da. Azken finean, helburua da euskararen garapen teknolinguistikoan lan egiten duten enpresek, unibertsitateek, agente teknologikoei eta gainerako interesdunek baliabide horiek erabili ahal izatea, hizkuntza-ereduak garatzeko, hobetzeko eta entrenatzeko, eta tresnak eta aplikazioak sortzeko.

Hauek dira bildu eta etiketatu beharreko hizkuntza-baliabideak:

- Idatzizko corpusak: corpus orokor elebakarrak (etiketatuak: morfologia, sintaxia, semantika, korreferentziak, koherentzia, rol semantikoa, izendun entitateak, denbora-adierazpenak, gertaerak eta abar), corpus eleaniztun paralelo edo konparagarriak (itzulpen-memoriak, testu elebidun edo eleaniztunak), arloetako corpus espezializatuak (hezkuntza, justizia, osasungintza, turismoa, administrazioa eta abar), eta, corpus historikoak edo ez-estandarrik.

Corpus ez-estandarrei dagokienez, aberasgarria litzateke euskalkietako formak edota, azken urteotan, sortzen ari den hizkuntza ez-formalean idatzitako corpus ugariak (sare sozialetakoak) biltzea eta tratatzea. Baina, Ekintza-planaren lehenengo fase honetan ez dira landuko, bestelako ezaugarriak dituztelako eta bestelako lanketa eta irizpideak behar dituztelako.

- Ahozko corpusak: egun, datu asko sortzen dira multimedia-formatuan plataforma eta gune desberdinetan (hedabideak, Internet ...) eta datu horiek eskuratzea eta etiketatzea funtsezkoa da daukaten informazio guztia modu egokian eta erosoan erabiltzeko eta ustiatzeko ahots-teknologietan: ahotsa testu bihurtu (ahots-ezagutza, *STT - Speech-to-text*) zein testua ahots bihurtu (ahots-sintesia, *TTS - Text-to-Speech*).

Ahozko corpusen artean, aipatzekoa da Mozilla Fundazioaren “Common Voice”⁶ proiektua, hainbat herrialde eta hizkuntzatan garatzen ari dena. Librezaile elkarteak sustatuta, euskarazko ahots-grabaketak biltzen ari dira Internet bidez, metadatuaz hornituak. Ondoren, grabazioak baieztatu egiten dira.

- Lexikoak, hiztegiak, entziklopediak, terminologia-bankuak, ezagutza-base lexiko-semantikoak (ontologiak): lexikoari dagokionez informazioa ematen dute, besteak beste, kategoría gramatikala baina baita hitzen arteko erlazio semantikoak ere (sinonimia, antonimia, hiperonimia / hiponimia).

4.1.3. Hizkuntza prozesatzeko lanabesak

Euskara hizkuntza-teknologietan integratzeko eta, hortaz, tresnak eta aplikazioak garatzeko, hizkuntza-baliabideak prozesatzeko eta erabilgarri izan daitezen prestatzeko, hau da ereduak sortzeko, zenbait lanabes behar dira, besteak beste, hauek:

- Analizatze eta etiketatze tresnak. Horiek izaten dira, askotan, konplexuenak:
 - Testuekin lotutako tresnak: analizatzaile morfologikoak, sintaktikoak, semantikoak (personak, erakundeak, lekuak... identifikatzeko), eta diskurtsoaren analizatzaileak. Bilaketak egiteko oso beharrezkoak dira. Horiek garatzeko teknikak aldatuz joan izan dira eta, gaur egun, gero eta gehiago erabiltzen ari da Adimen artifiziala.
 - Hizketarekin edo ahotsarekin lotutako tresnak: hizketa-ezagutzarako sistemak eta multimedia-materialak transkripzio aberastuaz etiketatze aukera ematen dutenak. Ezinbestekoak dira transkripzio automatikorako lanabesak lortzeko. Erraztu egiten dute multimedia-edukiak kudeatzea eta bilaketak egitea. Eredu generikoak sortu daitezke eta, ondoren, eremu espezifikotarako egokitu eta erabili.
- Datu pertsonalak eta entitateak anonimizatzeko tresnak. Corpusak eta itzulpen-memoriak garbitzeko ezinbestekoak dira.

4.1.4. Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea

Gaur egun, atzigarri dagoen informazio-masa (testuak, multimedia-produktuak...) handia da eta etengabe ari da sortzen, bai administrazio publikoan, bai hedabideetan, bai Interneten, oro har. Oso onuragarria litzatekeela baliabide horiek berrerabiltzeko baldintza egokiak izango balira.

Hori horrela, lehenengo urrats honetan, administrazio publikoan dauden, sortzen ari diren, eta datozen urteotan sortuko diren hizkuntza-baliabideak (testukoak nahiz ahozkoak) eta garatuko diren lanabesak eta tresnak plataforma batean jasotzeko urratsa eman nahi da. Horrela, “Datuen Bulegoa” plataforma sortuko da EJIeko zerbitzarietan. Bertan, jarriko dira proiektu desberdi-

6 <https://voice.mozilla.org/en/datasets>; <https://commonvoice.mozilla.org/eu>

netarako garatu eta garatuko diren corpusak, baliabideak, hizkuntza-ereduak eta tresnak, gero, sektoreko enpresek eta berrikuntza-zentroek kostu txikiagoarekin eta teknologia lehiakorrekin aplikazioak eta tresnak gara ditzaten. Arestian aipatu bezala, baliabideak banatzeko eta eskuratzeko protokoloa ezarri beharko da produktu bakoitzaren lizentziak errespetatuz.

Gainera, hizkuntza-baliabide horiek ustiatzeko eta hizkuntza-teknologietarako aprobetxateko, konputazio-ahalmen handia duten makinak behar direnez, abantaila izango da sektorearentzat datuak prozesatuta (etiketatuta) eta erabiltzeko prest eskuratzea.

4.1.5. Itzulpen automatikoa

Bistakoa da, egun, itzultzaile automatikoak ezinbesteko tresna bihurtu direla eguneroko jardueretarako (administrazio publikoa, eskola, enpresa, negozioak, Banka, Osasungintza, esparru sozio-ekonomikoa), baita aisialdirako ere (kirola, kultura...). Sare neuronaletan (Adimen artifiziala) integratu direnetik, jauzi handia gertatu da, itzultzaile automatikoen ematen duten kalitateari dagokionez. Tresna horiei esker, momentu honetan, askoz errazagoa da elebitasuna eta eleaniztasuna eguneroko erronka bihurtzea.

Euskara eskaintzen duten itzultzaile neuronal batzuk garatu dira azken urteotan. Eusko Jaurlaritzak ITZULI itzultzaile neuronal orokorra (<https://www.euskadi.eus/itzuli>) eta itzultzaile neuronal juridiko-administratiboa eskuragarri jarri ditu herritar guztion eskura; gainera, juridiko-administratiboaren kasuan, administrazio publikoek Euskalsarearen bidez erabil dezakete dokumentuak itzultzeko. Tresna hauek gero eta gehiago erabiltzen ari dira esparru profesionalean eta, pixkanaka-pixkanaka, ezinbesteko laguntza-tresna bihurtzen ari dira.

Hemendik aurrera, itzulpenaren kalitatea hobetzen jarraitzeaz gain, jarduera espezifikoetarako (Osasungintza, Justizia...) itzultzaile automatikoak garatu beharko dira, eremu horietako corpus ez eta terminologiak entrenatuta. Gainera, itzulpen gintza automatikoaren arloan, etorkizun hurbileko erronka ahots-ahots itzulpena izango da.

4.1.6. Hizketaren tratamendua: ahots-interfazeak

Adimen artifiziala ahots-teknologietan ere (sintesia eta ezagutza) erabiltzen ari da, eta, horri esker, kalitatea nabarmen handitu da. Azken urteotan, gero eta presenteago daude teknologia horiek egunero erabiltzen ditugun gailuetan, eskuko telefono, telebista eta bozgorailu adimentsuetan, besteak beste. Erabilera-erraztasuna eskaintzen dute eta erosotasun horretara ohitzen ari gara denok.

Hori dela eta, gero eta gehiago erabiltzen ari dira laguntzaile birtualak eta elkarrizketa-sistema automatikoak (chatbot-ak —testu bidezkoak ere—). Euskaraz ez dago, oraindik, horrelakorik, beraz, hori izango da lortu beharreko helburuetako bat.

Erronka da ahozko interfaze egokiak diseinatzea, eta ahots naturalak lortzea, espontaneoak, nor-tasuna eta emozioak adierazten dituztenak; horretarako, ahots anitzak eta adierazkorrak landu behar dira, tonu eta euskalki desberdinak kontuan izanda.



4.2 EKINTZA-PLANAREN ARLO ESTRATEGIKOAK

Hauek dira Ekintza-plan honetarako aintzakotzat hartutako arloak:

1. Administrazio elektronikoa
2. Hezkuntza
3. Helduen euskalduntzea
4. Osasungintza
5. Justizia
6. Turismoa
7. Kultura, hedabideak, humanitate digitalak, liburutegi digitalak...
8. Segurtasuna

Arlo horietan Hizkuntza-teknologiak integratzeko arrazoiak, lortu litezken helburuak eta garatu litezken atazak jaso dira segidan. Gainera, kasu bakoitzean, erabili beharreko teknologiak eta tresnak ere zerrendatu dira.

4.2.1 Administrazio elektronikoa

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>- Gobernantza Publiko eta Autogobernu Sailaren “ARDATZ Gobernantzaren, Berrikuntza Publikoaren eta Gobernu Digitalaren Plan Estrategikoa 2030” Planaren helburuen artean, hauek dira: Euskera laneko eta zerbitzurako hizkuntza izan dadin sustatzea; kudeaketa publiko aurreratuko sistemak izatea eta berrikuntza; unean-unean planteatzen diren erronkei egokitutako tresnak izatea; zerbitzu publiko modu proaktiboan eta pertsonalizatuan eskaintzea herritarrei; datuen gobernantza, eta informazioarena, segurua izatea”.</p> <p>- Administrazioen betekizunak:</p> <ul style="list-style-type: none"> · Gardentasuna: 19/2013 Legea · Komunikazioa eta parte hartzea: herritarrena eta eragileena. · Datuen babesa: Europako Parlamentuaren 2016/679 erregelamendua eta 3/2018 lege organikoa. · Inklusibitatea eta unibertsaltasuna: aniztasun funtzionala aintzat hartzea zerbitzu publikoak eskaintzeko orduan. 	<p><u>Administrazio elektronikoa eratzea</u>: administrazioek kudeatzen duten informazioa (testu- edo ahots-formatuan) euskarri elektronikoan maneiatzea, aurrerantzean kudeaketa ordenatua eta eraginkorra bideratu dadin.</p>	<ol style="list-style-type: none"> a. Leihatila bakarrean, eskaeren bideratze erdi-automatikoa. b. Espedienteen prozesamendu automatikoa. c. Gardentasuna. d. Akta automatikoak: transkribatu, laburtu eta egituratzea. e. Datuen anonimizazioa. f. Artxiboaren eta ondarearen digitalizazioa eta kontsulta. g. Inklusibitatea eta unibertsaltasuna: ahotsa eta testua ikuspegi unibertsalez lantzea. h. Informazioaren segimendua eta zaintza teknologikoa. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Ahots-ezagutza: transkripzioa • Itzultzaile automatikoa • Elkarrizketa-sistema automatikoak (chatbot-ak) • Dokumentuen bilaketa, erazketa, etiketatze eta sailkatze automatikoa • Zuzentzaile ortografikoa, gramatikala eta hizkuntza ez-sexistarena • Txantiloien araberako egituraketa automatikoak • Laburpen automatikoak • Datu pertsonalen anonimizazio automatikoa • Datu-baseen elikatze automatikoa • Testuen irakurketa errazteko laguntza-sistema automatikoa

4.2.2 Hezkuntza

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<ul style="list-style-type: none"> - 2018an Europako Batzordeak Hezkuntza Digitalaren Ekintza Plana (<i>Digital Education Action Plan</i>) zabaltzen hasi zen: teknologia digitalak (Adimen artifiziala, robotika eta hodei-konputazioa) era ezberdinetan aberasten ditu ikaskuntza-prozesuak eta, gainera, guztion eskura egon behar duten ikaste-aukerak eskaintzen ditu. - Heziberri 2020 marko pedagogikoan (236/2015 Dekretua EHAA) plana: Oinarrizko Hezkuntza amaitu duen ikasleak konpetentzia digitala eta mediatikoa izan behar du, konpetentzia digitalaren arloko esparru europarraren ildoan, egungo herritarrek eskatzen duten alfabetatze eta gaitzuntza funtzional osoa bermatzeko. - Hezkuntza inklusiboa. - ISEI Irakats-Sistema Ebaluatu eta Ikertzeko Erakundeak 2018an: “ikasleen euskalduntzea motel-tzen ari da”. - 236/2015 Dekretua, abenduaren 22koa, Oinarrizko Hezkuntzaren curriculum zehaztu eta Euskal Autonomia Erkidegoan ezartzen duena. 	<ul style="list-style-type: none"> - Testu-prozesamendurako eta ahots-tratamendurako teknika aurreratuak aplikatu eta integratzea: jakintza-arlo desberdinetako materialak automatikoki sortu eta aztertzea, ikasketa pertsonalizatua lantzea, ikasleen motibazioa sustatzea, gizarte-aldeak murriztea. - Sare sozialen erabilera hezkuntzaren ikuspuntu-tik lantzea: informazioa eta baliabideak bilatzea, pentsamendu kritikoa garatzea, ahalduntze kritikoak hezkuntza-prozesuen parte izatea. 	<p>[Hezkuntzaren maila guztietan]</p> <ul style="list-style-type: none"> a. Ikasle eta irakasleentzako tresnak garatu. b. Ikaskuntza-ingurune birtualak sortu. c. Hezkuntza-komunitatearen plataforma sortu (baliabide didaktikoak partekatzeko...) d. Hezkuntza arloko agente ezberdinen arteko lankidetzarako baliabideak prestatu. e. Jakintza-arlo desberdinetan lortzen den hizkuntza-maila lantzeko tresnak garatu. f. Datuen anonimizazioa bermatu. g. Sare sozial eta albiste atarietan agertzen diren albiste faltsuen detekzio automatikoa. h. Cyberbulling-a atzemateko tresnak garatu. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Laburpen automatikoak • Galdera-erantzunak • Idazten laguntzeko, idazkiak zuzentzeko, ikus-entzunezko materialetan etiketak jartzeko edota transkribatzeko eta kudeatzeko tresnak • Hezkuntza-corpusak bildu eta prozesatu • Ikasteko eta trebatzeko tresna espezializatuak • Sare sozialetako elkarrizketetan sentimenduen analisisa automatikoki egin

4.2.3 Helduen euskalduntzea

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<ul style="list-style-type: none"> - Testuinguru digitaleko irakaskuntza: teknologia-ren garapena eta kompetentzia digitalak garatzeko aukerak ez dira behar beste ustiatzen. - Egun euskaraz erabiltzen diren sistemak edo baliabideak (Ikasbil, BOGA edo ARIAN, esaterako) hobetu behar dira eta hizkuntza-baliabideak eta -teknologia integratu egin beharko lirateke PRIMERAN edo ARIAN bezalako ingurune digitaletan. - Ikaslearengan oinarrituriko autoikaskuntzako eta auto-ebaluazioa ahalbideratuko duten tresnak diseinatu behar dira. 	<p>[HABE + euskaltegiak + ikerketa-taldeak]</p> <ul style="list-style-type: none"> - Euskaltegiatako irakasleak hizkuntza-teknologietan trebatzea. - Euskaltegietan sortzen diren datuak eta materialak berrerabiltzea eta hizkuntza-baliabideak sortzea (corpusak, gidak...). - Hizkuntza-teknologia euskaltegietan erabilgarri izateko plana egitea: batetik, garatu diren tresnak moldatu eta erabilerraz jarri eta, bestetik, beharrezko tresna estrategikoak detektatu eta horiek diseinatu, datuetan oinarrituriko hizkuntza irakaskuntza sustatzeko eta ebaluazio-sistemak estandarren arabera egiteko eta errazteko. 	<ul style="list-style-type: none"> a. Idatzizko eta ahozko testu-bilduma corpus bat osatu eta etiketatu (atazak eta komunikazio-aldegiak). b. Ikas-materialak (dokumentazioa, bideoak, audioak) biltzen dituen gunea (bankua) sortu. c. Hizkuntza online ikasteko ingurune zabal eta erabilerraza garatu. d. Autoikaskuntzako tresnak eta ikas-materialak sortzen laguntzeko baliabideak prestatu. e. Testuak modu automatikoan aztertu. f. Corpus multimodaletan bilaketak egiteko eta datuak bistaratzeko tresna sendoak sortu. g. Ikus-entzunezko materialetan, ahotsaren ezagutza aplikatuta, azpigituluak automatikoki jarri. h. Testuak automatikoki sailkatu (A1-2/B1-2/C1-2) <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Zuzenketa automatikoa • Ahots-ezagutza: transkripzioa • Itzultzaile automatikoa • Elkarrizketa-sistema automatikoak (chatbot-ak) • Dokumentuen bilaketa, erauzketa, etiketatze eta sailkatze automatikoa • Zuzentzaile ortografikoa, gramatikala eta hizkuntza ez-sexistarena

4.2.4 Osasungintza

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>EAEn euskara eta gaztelania hizkuntzak koofizialak dira. Maiz osasun-langileekin komunikazioa euskaraz izaten den arren, haiek osasun-txostenak gaztelaniaz idazten dituzte. Osakidetzak sistema zentralizatu du eta osasun-txostenak langile guztiek ulertzea oso garrantzitsua da gaixoaren segurtasunerako.</p> <p>Osakidetzaren 2017-2020rako erronka eta proiektu estrategikoen artean bi ekintza zehatz ezarri ziren, dokumentazio eta historia klinikoan euskararen presentzia handitzeko:</p> <ul style="list-style-type: none"> - Euskara gero eta presenteago egitea dokumentazio klinikoan eta historia klinikoan. - Aurrera egitea historia kliniko bakar eta elebidun bateranzko bidean. <p>Osasun-langile profesionalak gero eta gehiago eskatzen dute historia klinikoan euskara erabili ahal izatea, baina profesionalen zati handi batek ez ditu bi hizkuntzak ezagutzen.</p> <p>Hizkuntza-teknologiek argia ekar dezakete, informazioa erauzteko teknikak garatzen dituztelako, hizkuntza eta testu mota anitzetarako.</p>	<p>Osakidetzan bi hizkuntza koofizialak arazorik gabe erabiltzea, ulermen arazoei teknologien bidez aurre eginez.</p> <p>Badaude osasun-alorreko gaztelaniazko testuak prozesatzeko tresna egokiak, baina euskararako egokitze- eta garapen-lan handiagoa behar da. Tresnak elebietan izan behar dira, pazienteen hizkuntza-eskubideak bermatu ahal izateko.</p> <p>Itzulpen automatikoa oso tresna lagungarria izan daiteke. Baina, itzultzaile automatiko espezifikoak garatu behar da eta, horretarako, osasun-txostenen corpus paralelo zabal eta osatua sortu behar da. Osasun-txosten elebakarrak ere oso erabilgarriak dira emaitzak hobetzeko, baina gutxi dago.</p> <p>Ahotsaren ezagutza eta prozesamendua ere ezinbestekoa da osasun-langile eta paziente gehiagotara iristeko.</p> <p>Eragin handiko esparrua da Osasungintza, azken finean gizarte osoak erabiltzen baititu osasun-zerbitzuak.</p>	<ol style="list-style-type: none"> a. Corpus paraleloak sortu (osasun-txostenak, itzulpen-memoriak) eta corpus elebakarrak ustiatu. b. Dokumentazio klinikoa prozesatu. c. Dokumentazio klinikoaren prozesamendurako kate (pipeline) ireki eta eleaniztun bat garatu. d. Osasun-txostenen idazketarako laguntza tresnak sortu (diktaketa). e. Osasun-txostenen itzulpen-sistema automatikoa prestatu. f. Baliabide terminologiko eta lexikoak sortu. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Dokumentuen bilaketa, erauzketa, etiketatze eta sailkatze automatikoa • Datu pertsonalen anonimizazio automatikoa • Testu-prozesamendurako tresnak • Ahotsaren teknologiak: sintesia eta ezagutza • Itzulpen automatikoa eta postedizio automatikoa • Zuzentzaileak: ortografikoa, gramatikala • Erreduzko txantiloien estandarizazio eta sorkuntza • Hiztegi elebidun eta eleanitzak • Terminologia erauzketa automatikoa, laburduren hedapena eta desanbiguzioa

4.2.5 Justizia

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>Dokumentu Judizialeko Zentroak urteak daramatza legegintzari eta jurisprudentziari buruzko informazioaren aukeraketa, antolaketa, tratamendua, zabalpeneak eta argitalpena sakonki lantzen.</p> <p>Zentroak garatutako baliabideen artean azpimarratzekoa da sei milioi dokumentuz osatutako jurisprudentziako corpusa. Karrera Judizialeko kideek baliabide telematikoen bitartez corpusa bere osotasunean eskuragarri dute. Publiko zabalari, berriz, corpusaren zati bat eskaintzen zaio, aurretik anonimizazio prozesu bat burutu eta gero. Horrez gain, tesauru juridiko bat garatu da, eta berau LEGIVOC proiektuan erabili da zuzenbide-terminologia nazional ezberdinen arteko interfaze semantiko bat eraikitzeko.</p> <p>Beraz, gaur egun, dokumentazio juridikoa eta jurisprudentzia biltzen dituzten gordailu handien kudeaketak hizkuntzaren tratamendu automatikorako teknika aurreratuak, ikasketa automatikoa eta analisi semantikoa barneratzen ditu dagoeneko. Horien bitartez, digitalizatutako dokumentuen gaineko informazio-erazketa, katalogazioa eta sailkapena modu homogeneo eta automatikoa burutzen da. Hala ere, prozesu hauen kostu handia dela eta, oraingoz, kide anitzeko organoetako epaiak besterik ez dira prozesatzen. Komenigarria litzateke, beraz, epai guztiak prozesatzea (kide bakarreko organoenak barne), baita administrazioan sortzen eta kudeatzen diren milioika dokumentu juridikoak ere.</p>	<p>Esparru juridikoan eraginkortasuna (bikoiztasunak ezabatzea) eta efikazia (martxan dauden produkzio-prozesuak hobetzea eta funtzionalitate eta gaitasun berriak gehitzea) hobetzen laguntzea. Testu-prozesamendurako, ahots-tratamendurako eta itzulpen automatikorako teknika aurreratuak aplikatu nahi dira existitzen diren dokumentazio juridikoaren gordailuetan eta hizkuntza koofizialetan. Epai- eta lege-dokumentazioaren prozesamenduak, testu-prozesamendurako eta itzulpen automatikoko tekniken bitartez, eragin handia izan dezake arlo judizialean, baita gizarte osoan ere, justiziaren marka transbertsalagatik.</p> <p>Horrez gain, justiziaren lan profesionalaren fase ezberdinetara iristeko, ahotsaren teknologia (ezagutza eta prozesamendua) erabiltzea ezinbestekoa da.</p>	<p>a. Corpus paraleloak lerrokatu (itzulpen-memoriak) eta corpus elebakarrak prozesatu.</p> <p>b. Epai- eta dokumentazio-prozesaketa prestatu.</p> <p>c. Justizia arloko testuen prozesamendurako kate (pipeline) ireki eta eleaniztun bat eratu.</p> <p>d. Justizia alorrerako itzulpen automatikoaren sistema bat garatu.</p> <p>e. Baliabide terminologiko eta lexikoak sortu.</p> <p>f. Bista judizialen digitalizazioa bultzatu.</p> <p>g. Pertsonen eta dokumentuen egiaztatze automatikoa garatu.</p> <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Dokumentuen bilaketa, erazketa, etiketatze eta sailkatze automatikoa • Datu pertsonalen anonimizazio automatikoa • Laburpen automatikoa • Analizatzaile morfosintaktiko eta semantikoak • Itzulpen automatikoa • Hiztegi espezializatuak • Terminologiaren erazketa automatikoa • Ahots-teknologia: transkripzioa, indexazioa, ahots-interfazeak, pertsonak identifikatzea, ziurtapena eta egiaztatzea (iruzurgileak eta intrusioa hautemateko teknikak —<i>spoofing</i> eta <i>antispoofing</i>—)

4.2.6 Turismoa

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>Turismoak komunikazio-beharrak du ezaugarri, eta, hortaz, hizkuntzak behar ditu. Turismoaren sektoreak, gainera, ahalmen ekonomiko izugarria du.</p> <p>Gaur egun, turisten eta erabiltzaileen iritzia jasotzea ezinbestekotzat jotzen da Interneten, batez ere sare sozialetan, eta, horretarako, teknologia behar dira.</p> <p>Turismo-informazioa hizkuntza askotan emateko, edukien kudeatzaileak behar dira. Itzulpen automatikoa tresna garrantzitsua da, tokian tokiko hizkuntza eta beste batzuk ere prozesatu daitezkeelako modu errazean. Testu-ahots eta ahots-testu bihur-gailuak ere garrantzitsuak dira arlo honetan, irisgarritasuna eta komunikazioa handitzen dutelako.</p> <p>Bestalde, Eusko Jaurlaritzak argitaratu zuen “Euskal Turismoaren Plan Estrategikoa 2020” izeneko txostenak, ekintza-ildoak eta ekimen estrategikoak definitzerakoan, apustu garbia egiten zuen Turismo alorrean hizkuntza-teknologia baliatzearen alde. Hiru ataletan jasotzen zuen berrikuntzaren estrategiari ekiteko modua: teknologia-berrikuntza, marka eta komunikazioa, eta hizkuntza-prestakuntza bultzatzea.</p>	<p>Bi aplikazio multzo zehatz aurreikus daitezke turismoaren arloan: museoen eta ondarearen gaineko balio erantsia aztertzeko, eta sare sozialetan turisten iritzi-erazketa egiteko. Biak ala biak ingurune eleaniztunaren agertokian. Turismo-eragileen web-guneak ere helburu inportantea izan daitezke.</p> <p>Tresna hauek turismo arloko eragileen esku (administrazioa, partzuergoak, teknologia-enpresak) hodei-zerbitzu moduan eskaintzea izan daiteke irtenbide egokiena. Euskarari garrantzi berezia eman beharko litzaioke, euskarazko ondareari balio erantsia eman ahal izateko. Bisitariak, edonondik datorrela ere, ohartu behar du Euskal Herrira etorri dela, eta euskararen herrian dagoela.</p> <p>Horrez gain, gidari-lanetarako ere hizkuntzarekin lotutako laguntzak gara litezke aplikazio moduan. Euskal Herriko informazio turistikoa zerbitzatzeko chatbot eleaniztun bat ere berrikuntza interesgarria litzateke.</p> <p>Hizkuntza-paisaian ere, seinalizazioa garrantzitsua da. Aireportuetan, geltokietan edo autobideetako mezuetan aplikazio egokiak izatea aurrerapauso nabarmena izango litzateke; azken batean, inguru eleaniztunean oso lagungarria baita mezu horiek norberak bere hizkuntzan jasotzea, modu irisgarri eta ulergarrian.</p>	<ol style="list-style-type: none"> Museoen eta bestelako ondarea modu eleaniztunean kudeatzeko ingurunea sortu. Turisten iritzia jakiteko (eta zerbitzuak hobetzeko) hodei-zerbitzu bat eratu. Chatbot eleaniztuna prestatu Euskal Herriko informazio turistikoaren gainean. Turismo-gidariei zuzendutako itzulpen-zerbitzu automatikoen azterketa eta bideraketa egin. Seinalizazioari begirako app eleaniztuna garatu. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> Eduki eleanitzak kudeatzeko tresnak Itzulpen automatikoa Ahots-teknologiak: sintesia eta ezagutza, ahoko interfazeak Elkarrizketa-sistema automatikoak (chatbot-ak) Iritzien inguruko informazioa prozesatzea (twitter) Webguneetan integratzeko api bat (hizkuntzak, geolokalizazioa, jaialdiak, ekitaldiak, festak... mapetan kokatuta) Laburpen automatikoa

4.2.7 Kultura, hedabideak, humanitate digitalak, liburutegi digitalak...

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>Kulturarekin lotuta, European garrantzi handia hartzen ari da Humanitate digitalen arloa (edo eHumanitateak), baita ondarea digitalizatzea eta hari beste balio erantsi bat ematea ere. Teknologiarik esker Humanitateetan paradigma-aldaketa bat gertatzen ari da (datuak).</p> <p>Digitalizazioarekin gero eta testu eskuragarri (gaur egungoak eta historikoak) gehiago ditugu, eta aho-tsa ere gero eta erabiliagoa da testuinguru digitaallean. Testuez eta hizketaren erregistroez gain, bestelako elementu batzuk ere ditugu kontuan hartzekoak: musika, bertsoak, argazkiak, ikus-entzuzkoak...</p> <p>Irratiak eta telebistak ere berebiziko garrantzia dute informazio-iturri gisa. Hedabideek erronka handia dute euskararen erabilerari begira.</p> <p>Gainera, kontsumitzeko modua asko aldatzen ari da digitalizazioarekin, batez ere gazteen artean.</p> <p>Kulturarekin lotutako informazio-iturri berriak ere agertzen ari dira digitalizazioarekin batera: blogak, podcast-ak eta sare sozialak. Ohiturak, iritziak, interesak, jakin-minak, eta abar detektatzeko tresnak garatzen ari dira.</p>	<ul style="list-style-type: none"> - Digitalizazioa, OCR zuzendua, bultzatzea: bilduma batzuk ez daude guztiz osatuta, beste batzuetan OCR egin gabe dago edo OCR prozesuaren akatsak zuzendu gabe. - Metadatuaren eskuragarritasuna bideratzea: estandarretan oinarrituta egon behar dute eta nazioarteko elkarrekinzarako prest. - Oinarritzko prozesaketa linguistikoa egitea: oinarritzko urratsak eman testuen gainean, eta emaitzak eskaini. Hau funtsezkoa da bilaketa eraginkorrek egin ahal izateko (lematizatuta, garaiaren arabera...) - Balio erantsia gehitzea: ikerketa metodo berriak erabili, teknologiaren arriskuak eta eragina neurtu, tresna linguistikoa aurreratuak aplikatu (metadatuaren aberasketa, adibidez), baita ere grafikoak eta mapak, besteak beste. <p>Prozesamendua aplikatzeko eremu interesgarriak dira hedabideak (hemeroteca/mediatekaren ustiapen aurreratuak) eta sare sozialak (soziolinguistikarekin lotutako aplikazioak). Euskarazko berriak eta euskal kulturaren edukiak, itzulpen automatikoaren bidez, euskaratik erdaretara itzuli (emaitza oso onak ari dira lortzen eu→es norabidean) eta, horrela, munduan zehar errazago hedatu.</p>	<p>[Artxiboak, liburutegiak, ondare-funtsak, hedabideak]</p> <ol style="list-style-type: none"> a. Biblioteka digitalen arloan, dokumentuak eta bestelako objektuak linguistikoki prozesatzeko kate bat eskaini, digitalizazioan laguntzeko. b. Humanitate Digitalak sustatu. c. Testu historikoei digitalki balio erantsia eman: facsimileaz gain, testuaren transkripzioa eta normalizazioa egin, eta bertako objektuak Wikipediarekin, edota liburutegiko beste lan batzuekin (edo itzulpenekin), estekatu. d. Hedabideei begira balio erantsiko tresnak garatu. e. Sare sozialen azterketa soziolinguistika egin: hiztunen portaera linguistikoa, hizkuntzaren eboluzioa, edota kode-alternantziaren fenomenoak. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • OCRa (optical character recognition); metadatuak osatzeko eta oinarritzko prozesaketa linguistikoa egiteko tresnak • Datuak eta corpusak, eta haien prozesaketa multimodal automatikoa • Dokumentuen multzokatze eta sailkatze automatikoak • Ahots-teknologia: testu-ahots bihurtzekoak eta transkripzioa • Itzultzaile automatikoa • Azpigituluen sorkuntza eta txertaketa • Lotutako berrien estekatzeko automatikoa

4.2.8 Segurtasuna

MOTIBAZIOA	HELBURUAK	ATAZAK, TEKNOLOGIAK ETA TRESNAK
<p>Ertzaintzan euskararen erabilera normalizatzeko II. plana, 2018-2022:</p> <p><i>“ertzainen euskara-gaitasuna handitzen jarraitu; euskararen erabilera sustatzen jarraitu, lankideekin eta herritarrekin; euskararen kudeaketa Ertzaintzako ohiko kudeaketa-sisteman txertatu; eta, informazio- eta komunikazio-estrategia eraginkorrak garatu.”</i></p> <p>Ertzaintzaren III. Plangintzaldiak 2027. urtera arteko iraunaldia du eta lau ardatz nagusi ditu:</p> <ol style="list-style-type: none"> 1. Ertzainen euskara-gaitasuna. 2. Euskararen erabilera herritarrekin eta lankideekin. 3. Euskararen kudeaketa. 4. Erabilera-planaren gaineko informazio eta komunikazioa. <p>Hizkuntza-teknologiek eskaintzen dituzten baliabideek eta tresnek ezinbestekoak dira “ertzainen euskara gaitasuna” hobetzeko eta, batez ere, idatzizko eta ahozko komunikazioan euskararen erabileran laguntzeko (zuzentzailea ortografikoa, hiztegi espezializatuak eta itzulpen automatikoa baliatuz).</p>	<p>Ertzaintzak herritarrekin dituen harremanetan hizkuntza-eskubideak bermatzea eta lehentasunezko unitate eta lan-eremuetan euskara ohiko lan-hizkuntza bilakatzea bultzatuko da. Horrek eragin handia izan dezake euskararen erabileran, bai Ertzaintzaren barnean, baita gizarte osoan ere.</p> <p>Ertzaintzan euskararen erabilera normalizatzeko III. Plangintzaldiaren helburu nagusia betetzeko, testu-prozesamendurako, ahots-tratamendurako eta itzulpen automatikorako teknikak aplika daitezke. Horrelako proiektuak martxan jartzeak eraginkortasuna hobetzen lagunduko luke. Lan profesionalaren fase ezberdinetara iristeko helburuarekin, testu-prozesamendua burutzeaz gain, hizketaren ezagutza eta prozesamendua burutzea ezinbestekoa da.</p> <p>Domeinu orokorreko baliabideak eta tresnak egokitu beharko dira domeinu zehatz honetara, honen ezaugarri eta beharrak kontuan hartuz.</p>	<ol style="list-style-type: none"> a. Dokumentazioaren prozesaketa. b. Dokumentazioaren prozesamendurako kate (pipeline) ireki eta eleaniztun bat garatu. c. Txostenen idazketarako laguntza-tresnak sortu. d. Txostenen itzulpen-automatikoa egin. e. Baliabide terminologiko eta lexikoak prestatu. f. Herritarrekin izandako harremanen digitalizazioa egin. <p>TEKNOLOGIAK ETA TRESNAK</p> <ul style="list-style-type: none"> • Anonimizazio prozesu estandar automatikoa (testuetatik informazio sentikorra kentzeko) • Txostenetako informazioaren eta terminologiaren erauzketa automatikoa • Prozesaketa linguistiko automatikoa: etiketatzea... • Ereduzko txantiloien sorkuntza eta estandarizazioa • Zuzentzaile ortografikoa eta gramatikala • Hiztegi eleanitz espezializatuak • Itzultzaile automatikoa eta postedizio automatikoa • Corpus paraleloak (itzulpen-memoriak) • Ahots-teknologiak: ahots-testu bihurketa, transkripzioen indexazioa eta bilaketa aurreratuak, ahots interfazearen inplementazioa

4.3 EUSKARAZKO HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANA, 2021-2024: ITSASARGI-PROIEKTUAK

Honela definitu ohi dira itsasargi-proiektuak: administrazio publikoek abiarazitako proiektuak, sektore estrategikoetan Hizkuntza-teknologiak aplikatzen dituztenak; teknologia horien ahalmena eta onurak erakusten dituzte, industria sortu eta bultzatzen dute, eta beste proiektu batzuetan berrerabil daitezkeen baliabideak sortzen dituzte. Gainera, etorkizunean, beste sektore batzuetan proiektuak bideratzeko ikaste-bidea ere izaten dira.

Itsasargi-proiektuek sinergia bilatzen dute Hizkuntza-teknologiak bultzatzeko martxan dauden gainerako neurriekin. Neurri horiek horizontalak eta transbertsalak dira (hizkuntza-azpiegitura orokorrak, hizkuntzaren ikuspuntutik interesa duen sektore publikoko informazioa berrerabiltzea, hizkuntza naturala prozesatzeko plataformak eta administrazio publikoetako itzulpen automatikoa), eta, sektore jakin batzuetan, balio-kate osoa hartzen duten eta produktu eta zerbitzu bokatuek ematen dituzten jarduketara zehatzak izaten dira.

Ekintza-plan honetarako estrategikotzat jotako arloetan, 2021-2024 legegintzaldian martxan dauden eta egitasmo diren proiektuak, Hizkuntza-teknologiak integratzen dituztenak, jaso dira.

Arestian esan moduan, ekintza-plan hau abiapuntua da, segida izango baitu. Hemen jasotako proiektuen jarraipena egingo da:

- Proiektuen bilakaera ezagutzeko,
- Proiektuek emandako emaitzen berri izateko,
- Proiektuak garatzeko sortutako eta erabilitako hizkuntza-baliabideak, ereduak, lanabesak eta tresnak berrerabiltzeko aukerak aztertzeo,
- Legegintzaldian zehar, arlo estrategiko horietan sortuko diren proiektu berriak planean integratzeko,
- Arlo estrategiko gehiagotara zabaltzeko ekintza-plana eta hizkuntza-teknologiak, eta
- Hizkuntza-teknologiak, proiektuak eta erabilgarri dauden tresnak ezagutarazteko administrazio publikoan, hezkuntza arloan, hedabideetan, lan-munduan eta gizarte osoan.

Zenbait kasutan ezin izan da fitxako informazio osoa osatu, oraindik ere finkatuta ez dagoelako. Informazio hori osatu eta eguneratuko da, proiektuen jarraipenarekin batera.

Atal inportantea da aurrekontua ezagutzea, horrek erraztuko baitu legegintzaldi osorako beharko den aurrekontua aurreikustea eta, hortaz, aurrera begira, prospekzioak egitea.

Guztira, 25 itsasargi-proiektu jaso dira fitxetan.





Proiektu-zk.: 1

IZENA	Herritarren arreta hobetu Adimen artifizialaren bidez
APLIKAZIO-ARLOA	ADMINISTRAZIO ELEKTRONIKOA
DESKRIBAPENA	<p>Proiektuaren ideia nagusia herritarrei erantzun automatikoak ematen dituen arreta emateko robota garatzea da (leihatila bakarra). Nukleoa osagai nagusia elikatu beharko litzateke besteak beste Zuzenean agenteek lehendik emandako erantzunekin eta <i>Maching Learning</i>-aren bidez sistema gai izan beharko litzateke aukera ezberdinen artean erantzun zuzena emateko.</p> <p>Sistema elebiduna izango litzateke. Testu bidez jasoko luke eskaera, eta erantzuna testu bidez itzuliko luke. Gainera geruza batean ahotsa testu bihur liteke; honela telefono bidez ere erabil liteke. Sistema erabilgarria izango litzateke ez bakarrik herritarrentzat, Zuzenean zerbitzuko langileentzat ere.</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: idatzizko corpusak - Hizkuntza prozesatzeko lanabesak: analizatzaile semantikoak, diskurtsoaren analizatzaileak - Hizketaren tratamendua: ahots-interfazeak, transkripzio automatikoa
HITZ GAKOAK	Leihatila bakarra, herritarren arreta-zerbitzu automatikoa
SAIL ARDURADUNA	GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Herritarrak Hartzeko eta Zerbitzu Digitaletako Zuzendaritza
GARATZAILEAK	Zuzenean, Web-alorra (Herritarrak Hartzeko eta Zerbitzu Digitaletako Zuzendaritza), EJIE
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	zehazteke

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.7 E-ADMINISTRAZIOA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas</i>



Proiektu-zk.: 2

IZENA	Datu pertsonalen anonimizaziorako aplikazioa
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	IVAP ari da lanean itzulpen-memoretako datu-pertsonalak anonimizatzeko eta ordezkatzeko tresna automatiko baten garapenean. Tresna hori garatzeko fasean dago eta behin anonimizazioa eginda Open Datako katalogoan publikatuko dira Itzulpen-memoria guztiak, edozeinek berrerabiltzeko aukera izan dezan. Datu hauek kanpoko edozein enpresa edo erabiltzailerri emango lioke aukera itzultzaile automatikoak garatzeko edo hobetzeko eta Hizkuntza-teknologiaren sektorea sustatzeko.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: idatzizko corpusak - Hizkuntza prozesatzeko lanabesak: analizatzaile morfosintaktikoak eta semantikoak - Itzulpen automatikoa
HITZ GAKOAK	Datu-babesa bermatu, sektore publikoaren hizkuntza-baliabideak berrerabili
SAIL ARDURADUNA	GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Herritarrak Hartzeko eta Zerbitzu Digmaletako Zuzendaritza, IVAP, EJE
GARATZAILEAK	EJIE, IVAP
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	75.000 € (15.000 + 60.000)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.5 SEGURTASUNA</p> <p>5.7 E-ADMINISTRAZIOA</p> <p>5.8 HEZKUNTZA</p> <p>5.11 E-OSASUNA</p> <p>5.12 E-JUSTIZIA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p> <p>5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p>- <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público</i></p>



Proiektu-zk.: 3

IZENA	Adimen artifizialaren sustapena
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Lan asko dago egiteko Adimen artifizialaren arloan eta, horretarako, zabaltasunean lan egin behar da. Enpresa-sarea sustatu behar da eta ideia berriak izatea ezinbestekoa da. Eusko Jaurlaritzako sailetan dauden aukerak miatzea eta, horrekin batera, enpresa teknologikoen pausoak ematea, hau da, biak, Administrazioa eta enpresak, parean ibiltzea izan daiteke bidea. Proiektuaren lan-lerro nagusiak hauek dira:</p> <ul style="list-style-type: none"> - Administrazioan aplikatu daitezken ideiak sustatzeko diru-laguntzak - Aukera tekniko berriak probatzea (Question&Answering - <i>Conversational Artificial Intelligence</i>): kontzeptu-probak <p>Hori dela eta, garrantzitsua izango dira bi hauek:</p> <ul style="list-style-type: none"> - Zaintza teknologikoa egitea - Sektorearen (merkatuaren) behaketa egitea (premia berriak atzemateko)
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak - Hizkuntza prozesatzeko lanabesak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua: ahots-interfazeak
HITZ GAKOAK	Gardentasuna, datu-babesa bermatu, leihatila bakarra, dokumentuen prozesamendua, aktak automatikoki jaso, artxiboaren digitalizazioa eta kontsulta, inklusibitatea eta unibertsaltasuna, informazioaren zaintza teknologikoa
SAIL ARDURADUNA	GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC)
GARATZAILEAK	Laguntza-deialdia: Deusto Seidor, Ibermática, LKS, Teknei, Legítimo, Serikat Teknologia probatzea (Q&A): EHUko Hitz Zentroa
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	200.000 € <ul style="list-style-type: none"> - Laguntza-deialdia: 90k+90k - Teknologia probatzea (Q&A): 20k

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK 5.14 TURISMOA ETA MERKATARITZA
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje II: Impulso de la Industria del Lenguaje</i> <ul style="list-style-type: none"> - Línea 1: Mejora de la visibilidad y la transferencia del sector - Línea II: Apoyo a la internacionalización y comercialización del sector <i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i>



Proiektu-zk.: 4

IZENA	Hizkuntza-teknologiaren zerbitzu orokor transbertsalak
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>2019an martxan jarri zen ITZULI itzultzaile neuronal eta gaur egun errotuta dago gizar-tean. Itzulpenaz aparte, hurrengo urratsak transkripzioa eta sintesia lantzea dira. Proiektuaren lan-lerro nagusiak hauek dira:</p> <ul style="list-style-type: none"> - Teknologia erostea erkidego osoan erabili ahal izateko, eta itzulpen-, transkripzio- eta sintesi-zerbitzuak martxan jartzea. - Zerbitzuen kudeaketa teknika antolatzea sailen erabileraren zein jardueraren bilakaeraren arabera.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza prozesatzeko lanabesak: diskurtsoaren analizatzaileak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua: transkripzio automatikoa, ahots-interfazeak
HITZ GAKOAK	Zerbitzu transbertsalak, dokumentuen prozesamendua, inklusibitatea eta unibertsaltasuna, informazioaren zaintza teknologikoa, sektore publikoaren hizkuntza-baliabideak berrerabiltzea
SAIL ARDURADUNA	GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC)
GARATZAILEAK	EJIE, VICOMTECH
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	<p>800.000€ (4 urte)</p> <ul style="list-style-type: none"> - Teknologia: 200k+200k+200k+200k <p>Oharra: 100.000 €, ITZULI itzultzaile neuronalaren ereduak hobetzea eta eguneratzea (ikus 17. fitxa)</p>

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK 5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas</i> <p><i>Eje IV: Proyectos piloto de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 5

IZENA	INGURA: Helduen euskalduntzerako ikas-plataforma integratua			
APLIKAZIO-ARLOA	HELDUEN EUSKALDUNTZEA			
DESKRIBAPENA	Helduen euskalduntzerako ikas-plataforma integratu bat sortzea, HEOCak (Helduen Euskalduntzearen Oinarrizko Curriculum) ezartzen dituen gidaleroei jarraikiz euskararen sei gaitasun-mailak lortu ahal izateko. Bere barruan izango ditu mailaz maila modu autonomoan eta komunitatean garatzeko ikastaro ikuskatu nahiz ikuskatu gabeak, ikas-irakaskuntza partehartzaile eta partekatzailea garatzeko komunikazio-tresnak, eta erabiltzaile-mota desberdinen (ikasle, irakasle, administrari...) arabera kudeaketa-eremuak.			
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizatze-tresnak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua: ahots-interfazeak, sintesia, transkripzio automatikoa 			
HITZ GAKOAK	Ikas-materialak, irakasleen prestakuntza, hizkuntza ikasteko ingurunea, materialak sortzeko baliabideak, testuak automatikoki aztertu, corpusetan bilaketak egiteko tresnak, ahots-ezagutzaren bidez bideoak azpigitulatu, chatbot-ak			
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA HABE, Hizkuntza Politikarako Sailburuordetza			
GARATZAILEAK	EJIE			
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021	<input checked="" type="checkbox"/> 2022	<input checked="" type="checkbox"/> 2023	<input checked="" type="checkbox"/> 2024
AURREKONTUA	2.000.000 € (2020-2026)			

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas</i> - <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la información del Sector Público.</i> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 6

IZENA	INGURA app-a: euskara ikasteko ikasgaien aplikazioa
APLIKAZIO-ARLOA	HELDUEN EUSKALDUNTZEA
DESKRIBAPENA	Euskara ikasteko ikasgaiak eskainiko dituen sakelako telefonoetarako aplikazio bat sortzea, edozein herritar euskarara hurbildu eta oinarrizko gaitasun komunikatiboa lor dezan. Gaur egun merkatuan dauden app-en antzekoa sortzea da helburua, HABEren oinarrizko curriculumaren (HEOC) gidalerroak aintzat hartuta. Aplikazioa independentea izan liteke edo INGURA plataformaren parte.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizatze-ko tresnak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua: sintesia, transkripzio automatikoa, ahots-interfazeak
HITZ GAKOAK	Ikas-materialak, irakasleen prestakuntza, hizkuntza ikasteko ingurunea, materialak sortzeko baliabideak, testuak automatikoki azertu, corpusetan bilaketak egiteko tresnak, ahots-ezagutzaren bidez bideoak azpigitulatu, chatbot-ak
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA HABE, Hizkuntza Politikarako Sailburuordetza
GARATZAILEAK	EJIE
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	INGURA plataformaren proiektuaren barruan dago.

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas</i> - <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la información del Sector Público.</i> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 7

IZENA	Testuak mailaren arabera lantzeko tresna baten garapena
APLIKAZIO-ARLOA	HELDUEN EUSKALDUNTZEA
DESKRIBAPENA	Proiektu hau HABE eta HITZ ikerketa-taldea ari dira garatzen elkarlanean, HPS eta UPV/EHuren 2020ko irailaren 15eko lankidetzeta-hitzarmenaren markoaren barnean (HABEk 2020an esleitutako ikerketa-beka baten lan-ildo bat da hau, eta 2021eko azaroan amaitu da). Ikasbil plataforman dauden testuak modu automatikoan mailakatzeko eta lantzeko tresna bat garatzea da proiektuaren helburua. Testu horietatik abiatuta egingo da lanketa automatikoa, bai mailakako konplexutasun-ezaugarri jakin batzuk webean bistaratzuz (adibidez, Tübingeneko unibertsitateak garatutako VIEW plugina), bai ikaslearen beharren arabera ariketak sortuz (adibidez, Tübingeneko unibertsitatearen FLAIR tresna).
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea; lexi-koak, terminologiak, ontologiak. - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua: sintesia, transkripzio automatikoa, ahots-interfazeak
HITZ GAKOAK	Ikerketa-programak, hizkuntza ikasteko ingurunea, materialak sortzeko baliabideak, testuak automatikoki aztertu, corpusetan bilaketak egiteko tresnak, nazioarteko erakundeekin elkarlana
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA Hizkuntza Politikarako Sailburuordetza EHU/UPV Bilboko Hezkuntza Fakultatea
GARATZAILEAK	HABE, HITZ Taldea (2021)
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	35.400 € (beka bidez, urtero)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i>



Proiektu-zk.: 8

IZENA	Egiaztatze gintzako proben corpusa osatzea
APLIKAZIO-ARLOA	HELDUEN EUSKALDUNTZEA
DESKRIBAPENA	Proiektu hau HABE eta HITZ ikerketa-taldea ari dira garatzen elkarlanean, HPS eta UPV/EHuren 2020ko irailaren 15eko lankidetzeta-hitzarmenaren markoaren barnean (HABEk 2020an esleitutako ikerketa-beka baten lan-ildo bat da hau, eta 2021eko azaroan amaitu da). HABEren egiaztatze gintza-prozesuko probekin corpus bat osatzea da proiektuaren helburua. Corpusa bi atal edo datu-basez osaturik egongo da: idazmen-proben datu-basea eta mintzamen-proben datu-basea. Corpusarentzako egitura egokiena zein izan litekeen aztertuko da, eta metadatu esanguratsuenen hautaketa bat ere egingo da. Proben transkripzioak nola lor daitezkeen ere aztertuko da.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Hizketaren tratamendua: transkripzio automatikoa, ahots-interfazeak
HITZ GAKOAK	Ikerketa-programak, testu-bilduma etiketatu, testuak automatikoki aztertu
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA Hizkuntza Politikarako Sailburuordetza EHU/UPV Bilboko Hezkuntza Fakultatea
GARATZAILEAK	HABE, HITZ Taldea (2021)
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	35.400 € (beka bidez, 2021, 2023), 70.000 € (2022)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i>



Proiektu-zk.: 9

IZENA	HOBE (HABEren Ordenagailu Bidezko EgiaztatzeGintza)			
APLIKAZIO-ARLOA	HELDUEN EUSKALDUNTZEA			
DESKRIBAPENA	HOBE sistema euskarazko komunikagaitasun-maila egiaztatzeko sortzen ari den tresna informatikoa da. Oraingoaz, azterketetako trebetasun-probak euskarri digitalean egiteko aukera ematen du, baina oinarri ezin hobea izan liteke Hizkuntza-teknologiak garatzeko testuak biltzeko (idatzizkoak zein ahozkoak), Adimen artifizialaren bidez datuak prozesatzeko (proben ekoizpenak automatikoki aztertzeko, adibidez) edo komunikazio digitalen funtzionaltasunak inplementatzeko (ahozko probak bideodei bidez egiteko, adibidez).			
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea; lexi-koak, terminologiak, ontologiak - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, transkripzio automatikoa - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Hizketaren tratamendua: ahots-interfazeak, ezagutza eta transkripzioa 			
HITZ GAKOAK	Hizkuntza ikasteko ingurunea, ahots-ezagutza, bideoak azpigitulatu, euskara prozesatzeko inguruneak sortu			
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA Hizkuntza Politikarako Sailburuordetza HABE			
GARATZAILEAK	HABE, EJIE			
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021	<input checked="" type="checkbox"/> 2022	<input checked="" type="checkbox"/> 2023	<input checked="" type="checkbox"/> 2024
AURREKONTUA	30.000 €, urtero (sistemaren garapena)			

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la información del Sector Público.</i> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 10

IZENA	ADITU-Osakidetza
APLIKAZIO-ARLOA	OSASUNGINTZA
DESKRIBAPENA	ADITU hizketa-ezagutzailea Osakidetzako Euskara Zerbitzu Korporatiboko lantaldeak erabiltzea, transkripzioak, azpituduluak eta itzulpen automatikoak egiteko. Batez ere, gaztelaniaz sortutako bideoen euskarazko bertsioa sortzeko erabiliko dira transkripzioak, azpituduluak eta itzulpen automatikoa; baina, horrez gain, batzarretan euskaraz egin nahi denean eta interpretazio-zerbitzua interpreteen bidez eman ezin denean erabiltzeko asmoa ere badago.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: ahozko corpusak - Hizkuntza prozesatzeko lanabesak - Itzulpen automatikoa - Hizketaren tratamendua: ahots-interfazeak, transkripzio automatikoa
HITZ GAKOAK	Osasun-zerbitzu digitalak, ahozko teknologiak, itzulpen automatikoa
SAIL ARDURADUNA	OSAKIDETZA Giza Baliabideen Garapen eta Prestakuntzako Zuzendariordeztza Euskara Zerbitzu Korporatiboa
GARATZAILEAK	Elhuyar Fundazioa
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	4.840 €, urtero

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.11 E-OSASUNA ETA ONGIZATEA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i> <i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i>



Proiektu-zk.: 11

IZENA	IDITE-Osakidetza
APLIKAZIO-ARLOA	OSASUNGINTZA
DESKRIBAPENA	IDITE lexiko-egiaztatzaile profesionala Osakidetzaren estilo-liburuaren gomendioen arabera pertsonalizatu nahi da, hasieran Euskara Zerbitzu Korporatiboarentzat eta, gero, erakunde osora zabaltzeko (bi fase). IDITEk erabiltzen duen teknologia garatuta dago, baina Osakidetzaren estilo-liburuaren gomendioetara egokitu gabe. IDITE tresna Osakidetzan erabiltzearen helburua hau da: Osakidetzako Zuzendaritza Nagusian (1. fasean) euskaraz sortutako edo euskarara itzulitako testuetan euskarazko lexikoaren erabilera zuzena ziurtatzea. Euskarazko zuzentzaile ortografikoak ez daude beti erabilgarri (adibidez, Wordfast PROrako), eta horrez gain, gerta daitezke ortografikoki zuzena izanda ere, termino bat ez egoki erabiltzea bere testuinguruan, eta horretan lagun dezake IDITEk.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: lexikoak, terminologiak, ontologiak - Hizkuntza prozesatzeko lanabesak: analizatzaile morfosintaktikoak eta semantikoak
HITZ GAKOAK	Osasun-zerbitzu digitalak, baliabide terminologiko eta lexikoak
SAIL ARDURADUNA	OSAKIDETZA Giza Baliabideen Zuzendaritza – Euskara Zerbitzu Korporatiboa
GARATZAILEAK	UZEI, Osakidetza
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	7.260 € (urtero)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	5.11 E-OSASUNA ETA ONGIZATEA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p>- <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la información del Sector Público.</i></p> <p><i>Eje IV: Proyectos faro de tecnologías del procesamiento del lenguaje natural de las Administraciones Públicas. Proyectos de aplicación de las tecnologías del lenguaje natural en servicios públicos estratégicos de gran impacto social</i></p>



Proiektu-zk.: 12

IZENA	Itzulbide
APLIKAZIO-ARLOA	OSASUNGINTZA
DESKRIBAPENA	Osakidetza testu klinikoen itzultzaile automatiko neuronala garatzen ari da; balioztatzefasean dago. Ondoren, Osakidetzako estazio klinikoetan ezarriko da, eta tresna garatzen jarraitzeko ildoak finkatuko dira (medikuntzako zuzentzailea, terminologia, euskarazko laburdurak...)
ARDATZAK	- Hizkuntza-baliabideak: testu nahiz ahozko materialak biltzea eta etiketatzea; terminologiak, ontologiak - Itzulpen automatikoa
HITZ GAKOAK	Dokumentazio klinikoaren prozesaketa, osasun-txostenak idazteko laguntza-tresna, osasun-txostenen itzulpen automatikoa, baliabide terminologikoak eta lexikoak
SAIL ARDURADUNA	OSAKIDETZA Giza Baliabideen Zuzendaritza – Euskara Zerbitzu Korporatiboa
GARATZAILEAK	EHUko Informatika Fakultateko HITZ Zentroa
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	90.000 € urtean (aurreikusia)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.11 E-OSASUNA ETA ONGIZATEA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i> <i>Eje II: Impulso de la Industria del Lenguaje</i> <i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i> <i>Eje IV: Proyectos faro de tecnologías del procesamiento del lenguaje natural de las Administraciones Públicas. Proyectos de aplicación de las tecnologías del lenguaje natural en servicios públicos estratégicos de gran impacto social</i>



Proiektu-zk.: 13

IZENA	Itzultzaile neuronal judiziala garatzea eta ezartzea
APLIKAZIO-ARLOA	JUSTIZIA BERRITZAILEA
DESKRIBAPENA	Itzultzaile neuronal juridiko administratiboaren esperientzia eta Hizkuntza Normalizazioko taldeak Justizia Administrazioan egindako itzulpen-memoriak aprobetxatu-ko dira euskara-gaztelania/gaztelania-euskara itzultzaile neuronal judiziala egiteko. Tresna hori Justizia Administrazioaren langileen, laguntzaileen eta kanpoko eragile juridikoen eskura jarriko da. Kudeatzaile judizial elektroniko berriaren eta haren aplikazio osagarrien funtzionaltasunei lotuko zaie.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu-corpusak (itzulpen-memoriak) biltzea eta etiketatzea; lexikoak, terminologiak, ontologiak - Hizkuntza prozesatzeko lanabesak: analizatzaile morfosintaktikoak eta semantikoak, anonimizatzeko tresnak - Itzulpen automatikoa
HITZ GAKOAK	Justizia digitala, itzulpen automatikoa
SAIL ARDURADUNA	BERDINTASUN, JUSTIZIA ETA GIZARTE POLITIKETAKO SAILA Justizia Digitalaren eta Azpiegituren Zuzendaritza
GARATZAILEAK	EJIE
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	zehazteke

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i> <i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas</i>



Proiektu-zk.: 14

IZENA	Bideokonferentziak testualizatzeko tresna garatzea eta kudeatzaile prozesal elektronikoan (KPE) ezartzea, modu integratuan
APLIKAZIO-ARLOA	JUSTIZIA BERRITZAILEA
DESKRIBAPENA	Bideo-grabazio batean (ahozko jarduketa judizialak eta grabatutako bista judizialak) jasotako ahotsaren testualizazio automatikoa egitea, bideo-grabazioaren une zehatzetara joateko aukera ematen duten bilaketak egiteko, eta, horrela, epaileen lana arintzea, epaiak egiteko orduan, edo fiskalena, kalifikazioak egiteko orduan. Eginda-ko testualizazioa Avantiusen KPE berriarekin integratu beharko da.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: transkripzio automatikoa; bilaketa automatikoak - Hizketaren tratamendua: ahots-interfazeak, ahots-ezagutza eta transkripzioa
HITZ GAKOAK	Justizia digitala, bista judizialen digitalizazioa
SAIL ARDURADUNA	BERDINTASUN, JUSTIZIA ETA GIZARTE POLITIKETAKO SAILA Justizia Digitalaren eta Azpiegituren Zuzendaritza
GARATZAILEAK	EJIE, Justizia Digitalaren eta Azpiegituren Zuzendaritzaren Garapen Zerbitzua
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	zehazteke

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	3.2 ADIMEN ARTIFIZIALA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i> <i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas</i>



Proiektu-zk.: 15

IZENA	Azpidazketa eta itzulpen automatikoa ikus-entzunezko edukietan
APLIKAZIO-ARLOA	KULTURA, HEDABIDEAK, HUMANITATE DIGITALAK, LIBURUTEGI DIGITALAK
DESKRIBAPENA	<p>Proiektua martxan dago 2019tik.</p> <p>Hiru helburu nagusi ditu:</p> <ol style="list-style-type: none"> 1.- Azpidazketa automatikoa lortzea ikus-entzunezko edukietan (euskaraz eta gaztelaniaz). 2.- Behin hori lortuta eta finduta, itzulpen automatikoak egitea euskaratik gaztelaniara eta gaztelaniatik euskarara. 3.- Edukien dokumentazioa aberastea eta edozein bilatzailearen bitartez eginiko ikus-entzunezko eduki digitalen bilaketak erraztea, ahots-ezagutzaren bitartez eginiko azpitolazioa aprobetxatuz.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: transkripzio automatikoa - Hizketaren tratamendua: ezagutza, transkripzioa - Itzulpen automatikoa
HITZ GAKOAK	Dokumentuen digitalazioa, ahots-ezagutza, ikus-entzunezkoak azpitolatu eta itzuli, bilaketa automatikoak, hedabideentzako balio erantsiko tresnak
SAIL ARDURADUNA	<p>EITBko Ingenieritza eta Esplotazioa Zuzendaritza</p> <p>VICOMTECH</p> <p>Ikerketa aplikatuko Institutua</p>
GARATZAILEAK	EITB, Vicomtech, Ikerketa aplikatuko Institutua
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	400.000 € (2021-2024) (2020: 124.000 €)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas</i></p>



Proiektu-zk.: 16

IZENA	Common Voice proiekturako ahots-corpora sortzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Mozilla Fundazioaren ekimen horretarako, audio-grabazioen datu-sorta libre sortuko da, kode irekikoa. Helburua: ahots-teknologiak garatzen dituen edozein enpresa edo entitatek hartu ahal izatea, hori erabilia produktu jakinak sortzeko. Hiru fasetan egingo da:</p> <ol style="list-style-type: none"> 1. fasea: 900.000 esaldiko testu-corpora osatu (datu pertsonalik gabea, formatu egokian eta lizentzia librean) eta Common Voice plataformara igo. 2. fasea: esaldi horien grabazioak egingo dira boluntarioen parte-hartzearen bidez. Mota guztietako ahotsak erabili beharko dira. 3. fasea: grabazioak, banan-banan, balioztatuko dira.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizatze-tresnak - Hizketaren tratamendua: ahots-interfazeak
HITZ GAKOAK	Dokumentuen prozesamendua, datu-babesa bermatu, inklusibitatea eta uniberteltasuna, hizkuntza-tresnak garatu, hizketaren tratamendua
SAIL ARDURADUNA	<p>KULTURA ETA HIZKUNTZA POLITIKA Euskara Sustatzeko Zuzendaritza</p> <p>Librezale.eus</p>
GARATZAILEAK	Vicomtech (1. fasea)
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	18.016,90 € (2021: 1. fasea)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK 5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje II: Impulso de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - Línea 1: Mejora de la visibilidad y la transferencia del sector - Línea II: Apoyo a la internacionalización y comercialización del sector <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p>



Proiektu-zk.: 17

IZENA	ITZULI itzultzaile automatiko neuronal (www.euskadi.eus/itzuli) hobetzen jarraitzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>2019an erabiltzaile guztien eskura jarritako itzultzaile automatiko neuronal hobetzen jarraituko da corpus berrieekin trebatuz:</p> <ul style="list-style-type: none"> - frantsesa↔euskara: Euskararen Erakunde Publikoaren corpusarekin - Gainerako hizkuntza-pareak: corpusak eskuratzen diren neurrian. <p>Funtzionalitate berriak sartzea aztertuko da: testu bat idaztearekin batera, haren itzulpena erakusten joatea...</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak biltzea eta etiketatzea - Hizkuntza prozesatzeko lanabesak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa
HITZ GAKOAK	Hizkuntza-tresnak, ikasle-irakasleentzako tresnak, hezkuntza-plataforma, ikasteko tresnak, humanitate digitalentzako zerbitzuak, hedabideentzako balio erantsiko tresnak, ikas-materialak, materialak sortzeko baliabideak, itzulpen automatikoa
SAIL ARDURADUNA	<p>GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Informazioaren eta Komunikazioaren Teknologien Zuzendaritza EJIE</p> <p>KULTURA ETA HIZKUNTZA POLITIKA - Euskara Sustatzeko Zuzendaritza</p>
GARATZAILEAK	EJIE
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	Ikus 4. fitxa

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK 5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i> - Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas</p> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 18

IZENA	ITZULI itzultzaile automatiko neuronal (www.euskadi.eus/itzuli) sustatzen eta zabaltzen jarraitzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Zabaltzen jarraitu behar da, ahalik eta erabiltzaile gehiagorangana irits dadin:</p> <ul style="list-style-type: none"> - www.euskadi.eus/itzuli webgunea: itzultzaile orokorra eta juridiko-administrati-boa - Trados eta Wordfasterako plugina: Administrazio publikoetako itzulpen-zerbitzuetan integratu da (IVAPeko IZO, hiru foru-aldundiak, hiru hiriburuetako udalak, Nafarroako Gobernua). - API bidezko integrazioa: Euskararen Erakunde Publikoa (Iparraldea) - Android eta iOSerako app-ak: etengabe eguneratzen dira <p>2022an integrazio berriak aurreikusi dira: Ertzaintza, Nafarroako Parlamentua ...</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza prozesatzeko lanabesak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa
HITZ GAKOAK	Hizkuntza-tresnak, laguntza-tresnak, itzulpen automatikoa
SAIL ARDURADUNA	<p>GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA</p> <p>Informazioaren eta Komunikazioaren Teknologien Zuzendaritza EJIE</p> <p>KULTURA ETA HIZKUNTZA POLITIKA - Euskara Sustatzeko Zuzendaritza</p>
GARATZAILEAK	EJIE
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	<p>APla: 8.000 € (2021+2022)</p> <p>App-ak: 20.000 € (2022)</p>

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.5 SEGURTASUNA</p> <p>5.7 E-ADMINISTRAZIOA</p> <p>5.8 HEZKUNTZA</p> <p>5.11 E-OSASUNA</p> <p>5.12 E-JUSTIZIA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p> <p>5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas</i> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 19

IZENA	ITZULI+ tresna-barrari funtzionalitate berriak gehitzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>2020an garatu eta Chromen edo Firefoxen, gehigarri moduan, instalatu eta gero, ITZULI+ tresna-barrak (www.euskadi.eus/itzuliplus) webgune batean nabigatzeaz batera, aukera hauek eskaintzen ditu: testu bat itzuli (ITZULIren 6 hizkuntza-pareak), hitzak 3 hiztegitan bilatu (euskara↔gaztelania, ingelesa, alemana), edota euskarazko Wikipedian bilatu.</p> <p>Bada, tresna-barrari funtzionalitate berriak gehituko zaizkio: Adimen artifizialaren bidez garatutako ahots sintetikoa, transkripzioa, lematizazioa, hiztegi gehiago eta testuak lantzeko aukerak.</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-balibideak eta tresnak - Itzulpen automatikoa - Hizketaren tratamendua: ahots-interfazeak, sintesia, ezagutza
HITZ GAKOAK	Hizkuntza-tresnak, ikasle-irakasleentzako tresnak, hezkuntza-plataforma, ikasteko tresnak, inklusibitatea eta unibertsaltasuna
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA Euskara Sustatzeko Zuzendaritza
GARATZAILEAK	EJIE, Teknei
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	32.000 € (2021) + 72.600 € (lematizatzaileak, sinonimoen hiztegia) (2021) 25.000 € (2022)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.13 HIZKUNTZAREN INDUSTRIAK 5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones públicas - Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público



Proiektu-zk.: 20

IZENA	Hiztegi-pluginak erabilia, testuetan hitzak markatu eta itzuli
APLIKAZIO-ARLOA	SEGURTASUNA
DESKRIBAPENA	Ofimatika-paketearen aplikazioetan, testu batean, hitz batean kokatzean, martaku eta beste hizkuntza bateko itzulpena lortzea, baita erabilera-adibideren batekin ere.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza prozesatzeko lanabesak - Hizkuntza-baliabideak eta tresnak
HITZ GAKOAK	Baliabide lexikoak, hizkuntza laguntza-tresnak
SAIL ARDURADUNA	SEGURTASUNA Telekomunikazioak eta Sistema informatikoak kudeatzeko Zuzendaritza (TSIKZ) – Garapenerako Laguntza Sistemak
GARATZAILEAK	zehazteke
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	zehazteke

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<ul style="list-style-type: none"> - 5.5 SEGURTASUNA - 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <ul style="list-style-type: none"> - <i>Línea 1: Plataformas de procesamiento del lenguaje natural y traducción automática en las Administraciones Públicas</i>



Proiektu-zk.: 21

IZENA	Ahots-mezuak transkribatzeko eta aldi berean itzultzeko sistema garatzea
APLIKAZIO-ARLOA	SEGURTASUNA
DESKRIBAPENA	Euskaraz ahozko mezu bat jasotzen denean (adibidez, ertzain-etxeetako harreran), mezu hori transkribatu, hau da testu bihurtu eta, aldi berean, gaztelaniara itzuli. Sistemak hizkuntza gehiago hartzea oso komenigarria litzateke.
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: ahozko eta idatzizko corpusak - Hizkuntza-tresnak: transkripzio automatikoa - Itzulpen automatikoa - Hizketaren tratamendua: ahozko interfazeak, sintesia, ezagutza
HITZ GAKOAK	Ahots-teknologiak, transkripzio automatikoa, itzulpen automatikoa
SAIL ARDURADUNA	SEGURTASUNA Telekomunikazioak eta Sistema informatikoak kudeatzeko Zuzendaritza Ertzaintzaren Informatika Atala (SIP)
GARATZAILEAK	zehazteke
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024 <input checked="" type="checkbox"/> zehazteke
AURREKONTUA	zehazteke

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	5.5 SEGURTASUNA 5.13 HIZKUNTZAREN INDUSTRIAK
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>Eje IV: Proyectos faro de tecnologías del procesamiento del lenguaje natural de las Administraciones Públicas. Proyectos de aplicación de las tecnologías del lenguaje natural en servicios públicos estratégicos de gran impacto social</i>



Proiektu-zk.: 22

IZENA	Administrazio publikoek sortzen dituzten hizkuntza-corpusak biltzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Adimen artifiziala erabiltzen duten teknologietan euskara integratzen jarraitzeko eta, ondorioz, gero eta arlo gehiagotan, erabiltzaileen eskura tresnak, aplikazioak eta lanabesak jartzeko, ezinbestekoak dira datu masiboak. Datu horiek corpusetatik lortzen dira (ahozkoak nahiz idatzizkoak; elebakarrak, elebidunak nahiz eleaniztunak) eta horien kalitatea kudeatzea garrantzitsua da.</p> <p>Administrazio publikoek egunero-egunero sortzen dituzte horrelako corpusak. Hori dela eta, asmoa da horiek biltzen joatea, hurrengo urrats batean, tratatzeko eta prestatzeko Adimen artifizialean erabiltzeko eta, aurrerantzean, partekatzeko.</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak bildu eta etiketatu - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizazio-rako tresnak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea
HITZ GAKOAK	Datu-babesa bermatu, informazioaren zaintza teknologikoa egin, sektore publikoaren hizkuntza-baliabideak berrerabiltzea
SAIL ARDURADUNA	KULTURA ETA HIZKUNTZA POLITIKA SAILA Euskara Sustatzeko Zuzendaritza
GARATZAILEAK	
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	Printzipioz, gasturik gabea

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA 5.5 SEGURTASUNA 5.7 E-ADMINISTRAZIOA 5.8 HEZKUNTZA 5.11 E-OSASUNA 5.12 E-JUSTIZIA 5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i> - <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público</i></p>



Proiektu-zk.: 23

IZENA	Adimen artifizialean erabiltzeko datuak tratatu eta prestatzea + azpiegiturak eskuratzea
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Adimen artifizialean oinarritutako lanabesak eta tresnak garatzeko hizkuntza-datuak behar dira. Datu horiek tratatu eta prestatu behar dira. Beraz, garapenaren arabera, jasota dauden corpusetako datuak tratatu eta prestatuko dira. Halaber, beharrezkoak direnean, datuak ad hoc sortzea ere aurreikusten da.</p> <p>Proiektuaren lan-lerro nagusiak hauek dira:</p> <ul style="list-style-type: none"> - datuen tipifikazioa (itzulpen-memoriak, ...) - datuen prestateta (anonimizazioa, etiketak, transkripzioak...), eta - azpiegitura teknikoak eskuratzea (txartel grafikoak-GPU)
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak bildu eta etiketatu - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizaziorako tresnak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Itzulpen automatikoa - Hizketaren tratamendua
HITZ GAKOAK	Datu-babesa bermatu, informazioaren zaintza teknologikoa egin, sektore publikoaren hizkuntza-baliabideak berrerabiltzea
SAIL ARDURADUNA	<p>GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC)</p> <p>KULTURA ETA HIZKUNTZA POLITIKA SAILA Euskara Sustatzeko Zuzendaritza</p>
GARATZAILEAK	Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC) EJIE, IVAP
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	200.000 €, urtero

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.5 SEGURTASUNA</p> <p>5.7 E-ADMINISTRAZIOA</p> <p>5.8 HEZKUNTZA</p> <p>5.11 E-OSASUNA</p> <p>5.12 E-JUSTIZIA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p> <p>5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p>- <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público</i></p> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 24

IZENA	Datuaren Bulegoa
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Gune bat sortuko da Adimen artifizialaren oinarria diren datuak (kasu honetan, Administrazio publikoenak) bertan jartzeko eta, horrela, berrerabiltzeko moduan eskaintzeko.</p> <p>Bulegoak Administrazioaren proiektuak aurrera eramateko sortuko diren datuak (hizkuntza-baliabideak, corpusak), ereduak eta tresnak jasoko ditu eta horiek berrerabiltzeko protokoloak landuko ditu. Protokolo horiek finkatuko dituzte datu horien berrerabiltzea eta zabaltzea edo partekatzea.</p> <p>Proiektuaren xedea euskarazko baliabideen <i>linked open data</i> ekosistema sortzeko lehen pausoak ematea da. Horretarako, <i>Linguistic Linked Open Data</i>-ren irizpideak aplikatzen saiatuko da.</p> <p>Behin zerbitzua martxan, beste entitate batzuen baliabideak ere bertan jasotzea aztertuko da.</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza-baliabideak: testu zein ahozko materialak bildu eta etiketatu - Hizkuntza prozesatzeko lanabesak: analizatzaileak, etiketatzaileak, anonimizaziorako tresnak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea
HITZ GAKOAK	Datu-babesa bermatu, hizkuntza-baliabideen zaintza teknologikoa egin, sektore publikoaren hizkuntza-baliabideak berrerabiltzea
SAIL ARDURADUNA	<p>GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC) Herritarrak Hartzeko eta Zerbitzu Digitaletako Zuzendaritza</p> <p>KULTURA ETA HIZKUNTZA POLITIKA SAILA Euskara Sustatzeko Zuzendaritza</p>
GARATZAILEAK	Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC) EJIE
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	100.000 € (2022)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.5 SEGURTASUNA</p> <p>5.7 E-ADMINISTRAZIOA</p> <p>5.8 HEZKUNTZA</p> <p>5.11 E-OSASUNA</p> <p>5.12 E-JUSTIZIA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p> <p>5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p>- <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público</i></p> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>



Proiektu-zk.: 25

IZENA	Ahozko erabilera neurtzeko tresna
APLIKAZIO-ARLOA	DENAK (ADMINISTRAZIO ELEKTRONIKOA, OSASUNGINTZA, JUSTIZIA, HEZKUNTZA, SEGURTASUNA...)
DESKRIBAPENA	<p>Aplikazioak ahozko jarduna neurtzen du, eta elkarrizketak hizkuntzaren arabera bereizi eta sailkatzen ditu.</p> <p>Neurtzen duen unitatea elkarrizketa da, eta elkarrizketa hori zein hizkuntzatan (euskaraz, gaztelaniaz edo beste batean) izan den jasotzen du.</p> <p>Mikrofono baten bidez eta aldez aurretik programatutako maiztasunarekin, aztertzeko ari den gunean izaten ari diren elkarrizketak jaso, transkribatu eta, gorde gabe, hizkuntzaren arabera bereizi eta kuantifikatzen ditu.</p>
ARDATZAK	<ul style="list-style-type: none"> - Hizkuntza prozesatzeko lanabesak - Hizkuntza-baliabideak eta tresnak eskuratzea eta berrerabiltzea - Hizketaren tratamendua: ahots-interfazeak
HITZ GAKOAK	Ahozko erabilera neurtu, hizketa-hizkuntza ezagutu, datu-babesa bermatu, informazioaren zaintza teknologikoa
SAIL ARDURADUNA	<p>GOBERNANTZA PUBLIKO ETA AUTOGOBERNUA</p> <p>Informazioaren eta Komunikazioaren Teknologien Zuzendaritza (IKTZ/DTIC) / EJIE</p> <p>KULTURA ETA HIZKUNTZA POLITIKA SAILA</p> <p>Euskara Sustatzeko Zuzendaritza</p>
GARATZAILEAK	EJIE
GARAPEN-URTEA(K)	<input checked="" type="checkbox"/> 2021 <input checked="" type="checkbox"/> 2022 <input checked="" type="checkbox"/> 2023 <input checked="" type="checkbox"/> 2024
AURREKONTUA	50.000 (2022 eta 2023)

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<p>3.2 ADIMEN ARTIFIZIALA</p> <p>5.5 SEGURTASUNA</p> <p>5.7 E-ADMINISTRAZIOA</p> <p>5.8 HEZKUNTZA</p> <p>5.11 E-OSASUNA</p> <p>5.12 E-JUSTIZIA</p> <p>5.13 HIZKUNTZAREN INDUSTRIAK</p> <p>5.14 TURISMOA ETA MERKATARITZA</p>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<p><i>Eje I: Apoyo al desarrollo de infraestructuras lingüísticas</i></p> <p><i>Eje III: La Administración Pública como impulsora de la Industria del Lenguaje</i></p> <p>- <i>Línea 2: Recursos lingüísticos de las Administraciones Públicas y política de Reutilización de la Información del Sector Público</i></p> <p><i>Eje IV: Proyectos faro de tecnologías de procesamiento del lenguaje natural de las Administraciones Públicas.</i></p>

5. EKINTZA-PLANAREN HASIERAKO AURREKONTU-AURREIKUSPENA

Aurrekontu hau ez da behin betikoa, zenbait proiekturen kasuetan zehaztu gabe dagoelako; beste kasu batzuetan, berriz, nahiz eta kopurua agertu, erabilitako urtez urteko kopuru zehatza jarri beharko da. Beraz, urtero eguneratuko da eta, gainera, proiektu berriak integratzea aurreikusten da. Hona aurreikusitako hasierako aurrekontua:

Zk.	PROIEKTUAREN IZENA	2021	2022	2023	2024	GUZTIRA
1	Herritarren arreta hobetu Adimen artifizialaren bidez					
2	Datu pertsonalen anonimizaziorako aplikazioa	15.000	60.000			75.000
3	Adimen artifizialaren sustapena		200.000			200.000
4	Hizkuntza-teknologiaren zerbitzu orokor transbertsalak	200.000	200.000	200.000	200.000	800.000
5	INGURA: Helduen euskaduntzerako ikas-plataforma integratua	333.000	333.000	333.000	333.000	1.332.000
6	INGURA app-a: euskara ikasteko ikasgaien aplikazioa					
7	Testuak mailaren arabera lantzeko tresna baten garapena	35.400	35.400	35.400		106.200
8	Egiaztatzegintzako proben corpusa osatzea	35.400	70.000	35.400		140.800
9	HOBE (HABEren Ordenagailu Bidezko Egiaztatzegintza)	30.000	30.000	30.000	30.000	120.000
10	ADITU-Osakidetza	4.840	4.840			9.680
11	IDITE-Osakidetza	7.260	7.260	7.260	7.260	29.040
12	Itzulbide	90.000	90.000	90.000	90.000	360.000
13	Itzultzaile neuronal judiziala garatzea eta ezartzea					
14	Bideokonferentziak testualizatzeko tresna garatzea eta kudeatzaile prozesal elektronikoan (KPE) ezartzea, modu integratuan					
15	Azpidazketa eta itzulpen automatikoa ikus-entzunezko edukietan	100.000	100.000	100.000	100.000	400.000
16	Common Voice proiekturako ahots-corpusa sortzea	18.016,90				18.017
17	ITZULI itzultzaile automatiko neuronal (www.euskadi.eus/itzuli) hobetzen jarraitzea					
18	ITZULI itzultzaile automatiko neuronal (www.euskadi.eus/itzuli) sustatzen eta zabaltzen jarraitzea	4.000	24.000			28.000
19	ITZULI+ tresna-barrari funtzionalitate berriak gehitzea	104.600	25.000			129.600
20	Hiztegi-pluginak erabilia, hitzak markatu eta itzuli testuetan					
21	Ahots-mezuak transkribatzeko eta aldi berean itzultzeko sistema garatzea					
22	Administrazio publikoek sortzen dituzten hizkuntza-corpusak biltzea					
23	Adimen artifizialean erabiltzeko datuak tratatu eta prestatzea + azpiegiturak eskuratzea	200.000	200.000	200.000	200.000	800.000
24	Datuaren Bulegoa	100.000				100.000
25	Ahozko erabilera neurtzeko tresna		25.000	25.000		50.000
	GUZTIRA	1.277.517	1.404.500	1.056.060	960.260	4.698.337

6. HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANAREN JARRAIPENA

Ekintza-plan honen abiapuntua diren itsasargi-proiektuen jarraipena egingo da eta, horrela, proiektu bakoitzaren azken emaitzaren berri izateaz gain, hura burutzeko garatu behar izan diren hizkuntza-baliabideak eta tratatu behar izan diren datuak, baita garatutako teknologiak eta produktuak ere, berrerabiltzeko aukera argiak ezagutuko dira.

Gainera, garrantzitsua izango da jakiteko, benetan, Eusko Jaurlaritzako sail desberdinetan (baita erakunde autonomiadunetan ere) zein proiektu garatzen ari den eta garatuko den, sinergiak aprobetxatzeko eta ahaleginak uztartzeko. Horretarako ere, urtez urte, proiektu desberdinetan egingo diren inbertsioak eta erabiliko den aurrekontua jasotzea datu argigarria izango da.

Ekintza-planean jasotako itsasargi-proiektuen jarraipenaz gain, 2021-2024 legegintzaldian zehar itsasargi-proiektu gehiago sortuko denez, ekintza-planera ekartzea izango da helburua. Izan ere, ekintza-plan hau ez da itxia, baizik eta irekia, etengabe osatu eta eguneratuko baita. Hori horrela, Eusko Jaurlaritzako sail eta erakunde autonomiadun guztiekin etengabeko harremana izango da.

Proiektuen gaineko informazioa urtean behin eguneratuko da azaroan. Dena den, Ekintza-plana egiteko osatu diren bi koordinazio-batzordeekin urtean bi bilera egitea aurreikusi da. Hauek dira koordinazio-batzordeak:

1. koordinazio-batzordea:

- KULTURA ETA HIZKUNTZA POLITIKA SAILA: Kultura Sailburuordetza; Hizkuntza Politikarako Sailburuordetza; HABE; EITB
- GOBERNANTZA PUBLIKOA ETA AUTOGOBERNUA SAILA: Informazioaren eta Komunikazioen Teknologien Zuzendaritza-EJIE; Herritarrak Hartzeko eta Zerbitzu Digmaletako Zuzendaritza; Gobernu Irekiaren Zuzendaritza; IVAP
- EKONOMIAREN GARAPEN, JASANGARRITASUN ETA INGURUMEN SAILA: Eraldaketa Digitalaren eta Ekintzailetzaren Zuzendaritza
- HEZKUNTZA SAILA: Ikaskuntzaren eta Hezkuntza Berrikuntzaren Zuzendaritza; Berritzegune Nagusia

2. koordinazio-batzordea:

- OSAKIDETZA: Giza Baliabideen Zuzendaritza - Euskara Zerbitzu Korporatiboa
- BERDINTASUNA, JUSTIZIA ETA GIZARTE POLITIKAK SAILA: Justizia Digitalaren eta Azpiegi-
turen Zuzendaritza
- TURISMO, MERKATARITZA ETA KONTSUMO SAILA: Turismo eta Merkataritza Sailburuor-
detza
- SEGURTASUN SAILA: Giza Baliabideen Zuzendaritza

Beste sail edo erakunde autonomiadun batzuen proiektu berriak etorriko balira Ekintza-planera, batzorde horietan integratuko lirateke.

Gainera, urtean behin bilera berezi bat egingo dute Ekintza-planaren kudeatzaileek eta Ekono-
miaren Garapen, Jasangarritasun eta Ingurumen Saileko Eraldaketa Digitalaren eta Ekintzaile-
tzaren Zuzendaritzak “Euskadiko Eraldaketa Digitalerako Estrategia 2025”-arekin uztartzeko eta
koordinatzeko.

Halaber, urtean behin, Ekintza-planaren jarraipenaren inguruan informazioa emango zaio Euska-
raren Aholku Batzordearen Ingurune Digitalean Euskara Sustatzeko Batzorde-atal Bereziari, hark
ere “Gomendioak 2021-2024” ekimenaren jarraipenean integra dezan.

Aurreikusi bezala, epe luzera begira planteatu da Hizkuntza-teknologiaren Ekintza-plana, beraz,
proiektu guztien jarraipen zehatza egiteak eta proiektu berriak txertatzeak asko lagunduko dute
euskara Hizkuntza-teknologietan eta, oro har, Adimen artifizialean, integratzen. Horri esker, as-
koz ere errazagoa izango da, aurrerantzean, premiak ezagutzea eta bide horretan aurrera egitea.
Gainera, jakin badakigu, teknologia horien garapena azkarra izango dela.

Hortaz, bi fitxa-eredu gehitu dira hemen:

- 6.1 2021-2024 Hizkuntza-teknologiaren Ekintza-plana: itsasargi-proiektuen
jarraipena.
- 6.2 2021-2024 Hizkuntza-teknologiaren Ekintza-plana: proiektu berriak.

Proiektu guztien jarraipena egiteko eta horrekin erlazionatutako informazioa eskura izateko, Hiz-
kuntza Politikarako Sailburuordetzak webgune bat sortuko du 2022an. Gainera, 2021-2024 al-
dian Ekintza-plan honen emaitzak txosten tekniko batean jasoko dira, ebaluazio eta balantze
moduan, hurrengo legegintzaldian hurrengo faseari ekin ahal izateko. Txosten tekniko hori argi-
taratu eta komunikatuko da.

6.1 2021-2024 HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANA: ITSASARGI-PROIEKTUEN JARRAIPENERAKO FITXA-EREDUA

Proiektu-zk.	<i>(Ekintza-plan honetakoa)</i>
Jarraipen-data	
IZENA	<i>(Ekintza-plan honetakoa)</i>
APLIKAZIO-ARLOA	<i>(Ekintza-plan honetakoa)</i>
SAIL ARDURADUNA	<i>(Ekintza-plan honetan dagokion fitxakoa)</i>
GARATZAILEAK	<i>(eguneratu, beharrezkoa izanez gero)</i>
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024 <i>(eguneratu, beharrezkoa izanez gero)</i>
AURREKONTUA	Aurreikusia: € <i>(eguneratu, beharrezkoa izanez gero)</i> Erreala: €
Proiektuaren egoera	<input type="checkbox"/> Hasieran <input type="checkbox"/> Erdian <input type="checkbox"/> Amaieran <input type="checkbox"/> AMAITUTA
Amaituta badago, proiektuaren azken emaitza	
Non eta nola dagoen erabilgarri	
Proiektua gauzatzeko garatutako hizkuntza-datuak (baliabideak) eta lanabesak	<input type="checkbox"/> Corpus idatziak (elebakarrak, elebidunak, eleaniztunak) <input type="checkbox"/> Ahozko corpusak (datu-base akustikoak, grabazioak, ...) <input type="checkbox"/> Lexikoak, terminologiak... <input type="checkbox"/> Analizatzaileak <input type="checkbox"/> Hizketarekin (ahotsarekin) lotutako lanabesak: transkripzioarako etiketatzaileak <input type="checkbox"/> Hizkuntza-ereduak (itzulpen-automatikoa, ahozko teknologiak) <input type="checkbox"/> Anonimizazioarako tresnak <input type="checkbox"/> Bestelakoak (zehaztu)
Proiekturako sortutako datuak berrerabilgarriak dira? Baiezkoan, kudeatzaileek dagokionarekin kudeatuko dituzte lizentziak.	<input type="checkbox"/> BAI <input type="checkbox"/> EZ
Gehitu produktuaren erabilerari buruzko datuak	
Gehitu informazio osagarria	

6.2 2021-2024 HIZKUNTZA-TEKNOLOGIEN EKINTZA-PLANA: PROIEKTU BERRIETARAKO FITXA-EREDUA

IZENA	
APLIKAZIO-ARLOA	<input type="checkbox"/> Administrazio elektronikoa <input type="checkbox"/> Hezkuntza <input type="checkbox"/> Helduen euskalduntzea <input type="checkbox"/> Osasungintza <input type="checkbox"/> Justizia <input type="checkbox"/> Turismoa <input type="checkbox"/> Kultura, hedabideak, humanitate digitalak, liburutegi digitalak... <input type="checkbox"/> Segurtasuna <input type="checkbox"/> Besterik (<i>zehaztu</i>)
DESKRIBAPENA	
ARDATZAK	<i>(Ikus Ekintza-planaren 4.1.2, 4.1.3, 4.1.4, 4.1.5 eta 4.1.6 atalak)</i>
HITZ GAKOAK	
SAIL ARDURADUNA	
GARATZAILEAK	
GARAPEN-URTEA(K)	<input type="checkbox"/> 2021 <input type="checkbox"/> 2022 <input type="checkbox"/> 2023 <input type="checkbox"/> 2024
AURREKONTUA	€

EUSKADIKO ERALDAKETA DIGITALERAKO ESTRATEGIA 2025	<i>(Ekintza-planaren kudeatzaileek betetzekoa)</i>
PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE	<i>(Ekintza-planaren kudeatzaileek betetzekoa)</i>

7. ERREFERENTZIAK

EUSKO JAURLARITZA Ekonomiaren Garapen, Jasangarritasun eta Ingurumen Saila (2021). *Euskadiko Eraldaketa Digitalerako Estrategia 2025*.

https://bideoak2.euskadi.eus/2021/03/30/news_67948/EEDE2025_Estrategia_EUS.pdf

https://bideoak2.euskadi.eus/2021/03/30/news_67948/ETDE2025_Estrategia_ES.pdf

EUSKO JAURLARITZA Gobernantza Publiko eta Autogobernu Saila

(2021). *ARDATZ Gobernantzaren, Berrikuntza Publikoaren eta Gobernu Digitalaren Plan Estrategikoa 2030*.

https://www.euskadi.eus/contenidos/informacion/ardatz_2030/eu_def/adjuntos/ardatz-2030-eu.pdf

https://www.euskadi.eus/contenidos/informacion/ardatz_2030/es_def/adjuntos/ardatz-2030-es.pdf

EUSKO JAURLARITZA Lehendakaritza

(2021). *Zientzia, Teknologia eta Berrikuntzako Plana Euskadi 2030*.

<https://www.ikerbasque.net/sites/default/files/files/ZTBP%20EUSKADI%202030%20.pdf>

https://www.innobasque.eus/uploads/attachment_files/pcti_euskadi_2030-60acacb14bf62.pdf

EUSKARAREN AHOLKU BATZORDEA. HIZKUNTZA POLITIKARAKO SAILBURUORDETTZA

(2021). *Euskararen ingurune digitala, Gomendioak 2021-2024*. Vitoria-Gasteiz:

Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia. ISBN978-84-457-3592-3

<https://www.euskadi.eus/eusko-jaurlaritza/hizkuntza-politika/> (Argitalpenak)

<https://www.euskadi.eus/gobierno-vasco/politica-linguistica/> (Publicaciones)

GOBIERNO DE ESPAÑA, Ministerio de Asuntos Económicos y Transformación Digital

(2015). *Plan de Impulso de las Tecnologías del Lenguaje*.

<https://plantl.mineco.gob.es/Paginas/index.aspx>

BARNE-TXOSTENA

EHU-KO HITZ ZENTROA. *Hizkuntza-teknologiak sustatzeko planaren txosten teknikoa*

(2020)



Garena izaten jarraitzeko
hizkuntza-teknologiak