

«Joera aldatzen ari da eta gero eta gehiago erabiltzen da ahotsa makinekin harremanetan jartzeko» | Euskal Herria

naiz.eus/es/actualidad/noticia/20190622/joera-aldatzen-ari-da-eta-gero-eta-gehiago-erabiltzen-da-ahotsa-makinekin-harremanetan-jartzeko

June 22,
2019

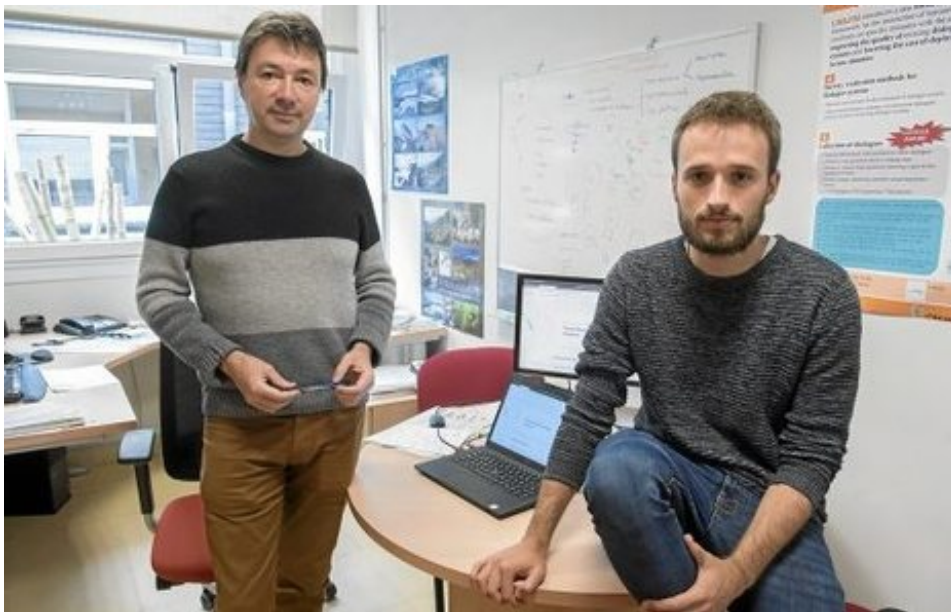
EUSKAL HERRIA

Eneko Agirre eta Jon Ander Campos

ixa taldeko kideak

Pertsonak makinekin harremantzeko modua aldatzen ari da. Gero eta gehiago, ahotsez gertatzen da, dialogo bidez. Teknologiari euskara irakasten ari direnak joera berriei begira daude. Hain zorrotz, oraintsu Google erakundearen sari garrantzitsu bat lortu dutela. Eneko Agirre irakasle eta ikerlariak eta Jon Ander Campos doktoregaiak azaldu digute saritutako proiektua.

«Joera aldatzen ari da eta gero eta gehiago erabiltzen da ahotsa makinekin harremanetan jartzeko»



«Izaten dira albiste beldurgarriak eta filma kezkarriak, baina teknologia oraintxe bertan oso sinplea da. Makinari galdera bat egitean, algoritmo baten bidez, erantzunaren hasiera eta bukaera identifikatu dezake»

«Googlek emandako sariak errekonozimendua eta prestigioa ekarri dizkigu. Eta, horrez gain, diruarekin jendea kontratatu ahal izango dugu proiektuarekin aurrera egiteko»

«Kalean ez badugu euskara erabiltzen, ez da entzuten, ez da existitzen. Teknologiarekin berdin, ez badaukagu nabigatzailea euskaraz, teknologia konpainiak ez dira konturatzen gu euskaldunak garela»

Google erakundearen ikerketa sarietako bat lortu dute. Mundu mailako sariketan, punta-puntako proiektuak lehiatzen dira. Guztira 158 proposamen saritu ditu Googlek –adimen artifizialaren esparruko 18 arlotako 910 hautagairen artean, 40 herrialdeetako 320 unibertsitatetik– eta UPV/EHUko Informatika Fakultateko Ixa taldeak aurkezturikoak lortu du sarietako bat, 80.000 dolarreko laguntza ekonomikoarekin.

Interneten dagoen informazioari buruz hitz egin ahal izateko elkarrizketa sistemak garatzea du helburu proiektuak. Finean, makinarekin harremantzeko modu berriak sortzen ari dira.

Nor dago saritutako proiektuaren atzean?

[Eneko Agirre] Saritu duten proiektuaren atzean Ixa taldea –<http://ixa.si.ehu.es/>– dago. Lengoaia naturalen prozesamenduan aitzindaria da Euskal Herrian. Duela hogeita hamar urte baino gehiago sortu zen eta hasierako motibazioa euskararen prozesamendua eta zuzentzaile ortografikoa izan ziren. Itzultzaile automatiko bat egiten hasi ginen. Urteak pasatu dira, baina ikerkuntza lerro hau ez da inoiz bukatzen. Azkenean, anbizioa da hizkuntza ulertzea eta hizkuntza ulertu arte egongo da zabalik ikerketa lerroa.

Teknologiari euskara irakasten ari zarete?

[E.A.] Hori da motibazioetako bat. Gertatzen dena da teknologiari euskara, ingelesa edo txinera irakasteak zailtasun bera duela; lantzen ditugun teknikak hizkuntza guztietarako balio baitute. Hala ere, egia da hizkuntza bakoitzak bere baliabideak behar dituela; hiztegiak, testu multzoak...

Eta, noski, hizkuntza handiek eta txikiek ez dituzte tresna berdinak eskura.

[E.A.] Tresnak eta baliabideak. Erabiltzen ditugun algoritmo askok adibideetatik ikasten dute. Orduan, euskarazko adibide kopuru handi bat emanda, ingeleserako kalitate bera lor daiteke. Finean, tresna berdina da. Arazoa da euskarazko adibide gutxiago daudela.

Google erakundearen ikerketa sarietako bat lortu duzue. Zertan datza zuen proiektua?

[Jon Ander Campos] Proiektu horren helburua da Interneten atzigarri dagoen informazioa dialogo bidez atzitzeko modua eskaintzea. Finean, gaur egun bilaketa asko egiten ditugu, telefonoaren bidez edo, baina joera aldatzen ari da eta gero eta gehiago erabiltzen dira ahotseko laguntzaile birtualak. Beraz, azken helburua izango litzateke bilaketa horiek dialogo bidez egitea.

Aldaketa ematen ari da? Ahotsak gero eta garrantzia handiagoa al dauka?

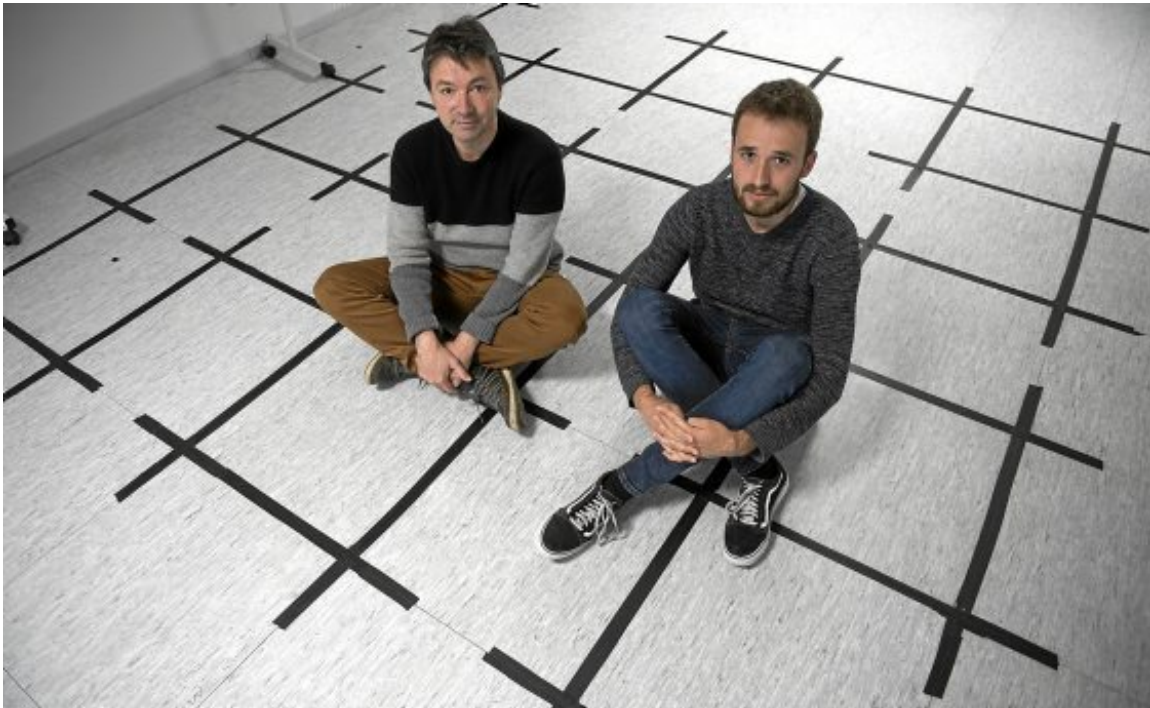
[E.A.] Bai. Orain arte asko erabili izan da bilaketa mota bat. Ordenagailuaren aurrean eserita zaude, zure bilaketa sartzen duzu eta dokumentu oso bat agertzen zaizu. Orduan, dokumentu hori guztia irakurri behar duzu interesatzen zaizuna topatzeko. Gertatzen

dena da beti ez gaudela eserita eta tekleatzeko moduan, gero eta gehiago konektatzen baikara Internetera telefono mugikorretik. Orain edozein egoeratan gaudela egiten dugu kontsulta.

Gainera, orain badugu lehen ez genuen gaitasun bat: dokumentuan dauden pasarteak errekonozitu ditzakegu eta galderak zuzenago erantzun. Adibidez, lehen bilatzailean galdetzen bazenuen zenbat biztanle dituen Donostiak, Donostiari buruzko dokumentu luze batera eramaten zintuen, eta hor datua topatu behar zenuen. Baina, orain, aldiz, galdera horri zuzenean erantzuten dio bilatzaileak berak. Kasu batzuetan, teknologia gai da dokumentua bilatu eta bertatik erantzuna ateratzeko.

Beraz, kontsulta horiek ahoz egingo dira gero eta gehiago.

[J.A.C.] Bai. Azken finean pertsonen artean interaktuatzeko dagoen modu naturala ahotsa den bezala, helburua da makinarekin interaktuatzeko modua ere ahotsa bihurtzea. Erosoagoa da.



Robotikarekin lotura duela dirudi. Alor horretan ere pertsonen eta teknologiaren arteko harreman ahalik eta naturalena bilatzen da.

[E.A.] Bai, neurri batean lotuta daude. Makinarekin komunikatzeko modu berri bat bilatzen ari gara. Esparru hau oso zabala da. Gauza bat da robot batekin komunikatzea. Adibidez, etxea garbitzen duen gailuarekin hitz egin ahal izango genuke eta agindu zehatzak eman. Eta beste gauza bat da sukaldeko errejeta baten inguruko zalantza bat galdetzea. Kasu horretan, dokumentu luzeago batetik aterako du erantzuna makinak. Gaitasun desberdinak dira. Googleren ikerketa sarietan aukeratu duten proiektua gehiago doa bigarren horretatik.

Kontsultak ahoz egin ahal izatea erosoagoa da. Baina pentsatzekoa da egoera jakinetan oso baliagarria izan daitekeela.

[E.A.] Bai. Adibidez irudika dezakegu pertsona bat sukaldean, eskuak zikinduta, eta errezetaren inguruko galdera bat egiten telefonoari. Edo norbaiti deitzeko agindua ematen ahotsez. Baina, era berean, ahotsaren bidez makinarekin gauzak egin ahal izateak bide berri asko irekitzen ditu, zaintzaren alorrean kasu. Gaitasun teknologiko handirik ez duen adin nagusiko jendeak ahotsez emango dio agindua makinari; edo berdin bere gaitasunak mugatuak dituen pertsonak.

Googleren sariak zer dakar?

[E.A.] Batetik, errekonozimendua eta prestigioa. Kalitatezko ikerkuntza errekonozitzen duen kanpoko sari bat da, finean. Eta, horrez gain, diruarekin jendea kontratatu ahal izango dugu proiektuarekin aurrera egiteko.

Zeintzuk izango dira proiektuaren hurrengo pausoak?

[E.A.] Proiektuaren hasieran gaude oraindik. Googlek saria eman aurretik proiektu europar bat genuen alor horretan. Orain, sariak proiektua indartu egiten du. Dagoeneko diseinatuta genuen zein datu jaso nahi genituen eta zer den makinak ikasi behar duena. Orain, askoz datu gehiago jaso behar ditugu eta algoritmo ahaltsuagoak erabili horretarako.

Eta deialdi publiko bat egin duzue horretarako.

[E.A.] Bai, bi kontu daude. Lehen esan dudan bezala, sistema hauek ikasten dutena baliagarria da edozein hizkuntzarako. Ikerkuntzaren munduan, nagusiki ingelesa da erabiltzen dugun hizkuntza algoritmo hoberena nori duen konparatzeko. Finean, gure Olinpiadak ingelesez egiten dira. Horretarako elkarriketa asko jaso behar ditugu eta dagoeneko badago plataforma bat horiek jasotzeko. Hori erraza da.

Baina Ixa taldearen helburua tresna teknologikoak euskaraz eskaintzea denez, deialdi irekia egin dugu euskarazko adibide ahalik gehien lortzeko eta ingeleserako garatu dugun teknologia hori euskarara aplikatu ahal izateko.

Nola josten dira algoritmoak osatzeko dialogo horiek?

[J.A.C.] Bi pertsonen arteko interakziotik. Pertsona batek dokumentu osoa dauka eta besteak dokumentu horri buruzko galdera zehatzak egiten dizkio. Gaur egun badaude "ohiko galderak" edo deitzen direnak web orrialde askotan. Bada proiektuaren helburuetako bat da "ohiko galderak" automatikoki eta dialogo bidez erantzuteko modua eskaintzea. Bi pertsonen arteko interakzio horietatik dialogo adibide mordoa lortzen dugu. Google erakundeak emandako sariari esker adibide gehiago lortu ahal izango ditugu -dialogo bakoitzak prezio bat daukalako-. Eta euskarazkoak ere bildu nahi ditugu.

Euskarazko dialogoak sortzeko boluntario bila zabiltzate. Zer egin behar da?

[J.A.C.] Bi pertsonen artean egin behar dute ariketa. Batek euskarazko Wikipediako artikulu bat izango du eskura eta besteak galderak egingo dizkio ahalik eta informazio

gehien lortzeko. Besteak erantzunak testutik hartuko ditu. Helburua da galdera bat hurrengoarekin lotuta egotea eta dialogo kutsua izatea. Ahalik eta euskarazko adibide gehien lortu nahi ditugu.

Idatziz egin beharko da ariketa, ezta?

[J.A.C.] Bai, ariketa idatziz izango da, ahotsaren kalitatearekin gorabeherarik ez izateko hobe da horrela.

Zein fasetan dago bilketa?

[E.A.] Orain boluntarioak jasotzen ari gara. Behin boluntarioak bilduta, beraien arteko elkarrizketan jarriko ditugu teklatu bidez. Norberak bere etxetik egin ahal izango du minutu gutxi batzuetan. Eta nahi adina elkarrizketa egin ahal izango ditu. Gainera, parte hartzaileen artean sari bat zozkatuko dugu. Milaka elkarrizketa behar ditugu. Jasotzen ditugun datuei esker espero dugu euskaraz beste aplikazio bat egotea, ingelesaren parekoa.

Hala ere, akaso lortuko dugu ingelesezko elkarrizketetatik ikasitakoa euskarara aplikatzea. Ondorioz, nahiz eta euskaraz adibide gutxiago izan, beharbada lortuko dugu ingelesezko kalitate bera elkarrizketa gutxiagorekin. Finean, pertsonok munduari buruz dakiguna, hizkuntza bat ikasten dugunean, hizkuntza horretara eramaten dugu.

Makina gero eta pertsonaren antzekoagoa da, beraz?

[E.A.] Galdera sinple horiek erantzuteko bai. Izaten dira albiste beldurgarriak eta filma kezagarriak, baina teknologia oraintxe bertan oso sinplea da. Finean, makinari galdera bat egitean, algoritmo baten bidez aukera izan dezake erantzunaren hasiera eta bukaera identifikatzeko. Kito. Hortik abiatuta norbaitek pentsatzen baldin badu makina adimentsua dela eta bere iritzi propioak dituela, oso oker dabil. Hori garbi eduki beharko genuke. Normalean, ikerlariak garenok barregura sortzen zaigu zenbait iragarpen katastrofiko entzutean. Aitzurra asmatu zenean norbaitek munduan aitzurrarekin zulo ikaragarri bat egin eta mundua aitzurkadaka txikituko zela esatea bezalakoa da. Oraintxe bertan aitzurraren mailan gaude.



Ikerlariak garbi daukazue teknologia laguntzeko dela, gure bizitza errazteko. Baina kontrako norabidean ere erabil daiteke.

[E.A.] Bai, eta horregatik da garrantzitsua unibertsitateetan eta ikerkuntza zentro publikoetan puntako gaietan jendea aritzea, erabilera ona eta txarra bereizteko gai izateko. Finean, ikerkuntza horiek zentro pribatuetan bakarrik egiten baldin badira, gizartea babesik gabe geratzen da, ez baitauka bere alde egingo duen inor. Horregatik oso garrantzitsua da ikerkuntza egitea, ez bakarrik aurrerapenak egiteko, baita erabilera txarra ematen denean horren aurrean erreakzionatzeko gaitasuna izateko ere.

Berez unibertsitateak zer interes dauka ikerkuntzan?

[E.A.] Zientzia zientziagatik, gizarteak aurrera egiteko. Eta gure kasuan, Ixa taldean eta Hitz ikerkuntza zentroan, euskara da zutabeetako bat. Gure helburuetako bat da euskara ez geratzea beste hizkuntzen atzean.

Nola kokatzen da euskara bezalako hizkuntza txiki bat teknologiaren mundu erraldoi horretan?

[E.A.] Hizkuntza txikiek beti handiek baino babes gutxiago daukate, baita teknologiaren

alor honetan ere. Berri ona da gaur egun, hizkuntzen teknologia hainbeste aurreratu denez, software industria erraldoien –Google, Microsoft, Apple...– eskuetan ere gero eta hizkuntza gehiago estaltzen direla. Adibidez, Googlek aspalditik dauzka gauza asko euskaratuta, nahiz eta hizkuntza txikia izan. Euskarak orain arteko indarrarekin edo gehiagorekin jarraitzen badu, ez dauka etorkizun txarra.

Horretarako erabiltzaileak euskararen aldeko hautua egin behar du, ezta?

[E.A.] Bai, hori da. Badaude mugikorra euskaraz edukitzeko deia egiten duten kanpainak eta garrantzitsua da. Kanean bezala gertatzen da: ez badugu euskara erabiltzen, ez da entzuten, ez da existitzen. Teknologiarekin berdin, ez badaukagu nabigatzailea euskaraz, teknologia konpainiak ez dira konturatzen gu euskaldunak garela. Gaztelerazko edo frantseseko erabiltzaileen artean agertuko gara.

Ikerkuntzan zaila da epeak zehaztea, baina zein da datozen urtetarako zuen plana?

[E.A.] Proiektu honek badu epe bat; adibidez, Jon Anderrek [Campos] tesia aurkeztu beharko du hemendik hiru-lau urtera eta ordurako gauza dezente egingo ditugu. Googleren proiektua, berriz, urtebetekoa da. Badakigu helburuak beteko ditugula baina asko geratuko da egiteko. Espero duguna da ikerketa lerro horretan datozen lau urtetan ahalik eta emaitza onenak lortzea. Finantzazioa dagoen bitartean eta arazoak erakargarriak diren bitartean, aurrera egiten da ikerkuntzan.