

ALBISTEAK
(/EU/CAMPUSA
/ALBISTEAK)
CATHEDRA
(/EU/CAMPUSA
/CATHEDRA)
KRONIKA
(/EU/CAMPUSA
/KRONIKA)
ERREPORTAJEAK
(/EU/CAMPUSA
/ERREPORTAJEAK)
IKERKETA
(/EU/CAMPUSA
/IKERKETA)

BASQUE | ESPAÑOL
([HTTPS://WWW.EHU.EUS:443](https://www.ehu.eus/443)
/EU/CAMPUSA
/EDUKIA?P_P_ID=82&

P_P_LIFECYCLE=1&
P_P_STATE=NORMAL&
P_P_MODE=VIEW&
P_P_COL_ID=COLUMN-
1&P_P_COL_POS=1&
P_P_COL_COUNT=4&
_82_STRUTS_ACTION=%2FLANGU
_82_REDIRECT=%2FEU%2F-%2
AUTOMATICA-
SIN-DICCIONARIO&
_82_LANGUAGEID=ES_ES |
ENGLISH

([HTTPS://WWW.EHU.EUS:443](https://www.ehu.eus/443)
/EU/CAMPUSA
/EDUKIA?P_P_ID=82&
P_P_LIFECYCLE=1&
P_P_STATE=NORMAL&
P_P_MODE=VIEW&
P_P_COL_ID=COLUMN-
1&P_P_COL_POS=1&
P_P_COL_COUNT=4&

_82_STRUTS_ACTION=%2FLANGU
_82_REDIRECT=%2FEU%2F-%2
AUTOMATICA-
SIN-DICCIONARIO&
_82_LANGUAGEID=EN_GB)

(/eu/campusatrusts/permissions
(/rss-noindex/UPVEHU
(https://www.ehu.eus/443)
(/asset-published/BUJMS)
CAMPUSATrusts/4431&fr
(/OIQO/rss?p_p_cacheability=cacheLevelFull)
(/eu/campusacampusari-
buruz)KRONIKAZIO

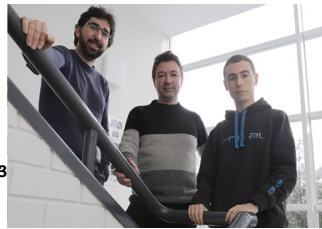
BULEGOA (/eu/komunikazio-
bulegoa) | NEWSLETTER
(/eu/campusa/newsletter) |
LEGE OHARRA
(/eu/campusa/lege-oharra) |
KONTAKTUA (/eu/campusa
/kontaktua)

< ATZERA

Itzulpen automatikoa, hiztegirik gabe

UPV/EHUko Informatika Fakultateko IXA taldearen ikerlan batek ikerketa ildo berri bat ireki du hizkuntzaren prozesamenduari esparruan

IKERKETA 2018/01/31



Ikasketa automatikoa, gehienetan, gizakiak gainbegiratzen du, eta itzulpen automatikoa ez da salbuespen bat: ordenagailuak pertsonak egindako milioika itzulpen erabiltzen ditu patroiak atera eta testu berria itzultzen ikasteko. Hurbilpen horrek ondo funtzionatzen du ingelesa eta frantsesa bezalako hizkuntza bikoteekin, euren arteko itzulpen ugari baitaude. Hizkuntza bikote gehienek, baina, askoz baliabide urriagoak dituzte, eta hori arazo handi bat da alemaniera-errusiera edo euskara-ingelesa bezalako hizkuntza bikoteetan. Testuinguru horretan, Universidad del País Vasco/ Euskal Unibertsitateko (UPV/EHU) Informatika Fakultateko IXA taldeko Mikel Artetxe, Eneko Agirre eta Gorka Labaka ikertzaileek ikasketa ez-gainbegiratu oinarritutako itzulpen automatikoko sistema bat garatu dute. «Imajinatu pertsona bati txineraz idatzitako hainbat liburu ematen dizkiozula, eta arabieraz idatzitako beste hainbeste liburu (elkarren artean ezberdinak), txineratik arabierara itzultzen ikas dezan. Hasiera batean, gizaki batentzat ezinezkoa dirudi. Baina guk frogatu dugu ordenagailu batek lan hori egiten ikas dezakeela», azaldu du Mikel Artetxek, hizkuntzaren prozesamenduari eta ikasketa automatikoari buruzko tesia egiten ari denak.

UPV/EHUko ikertzaileek proposatutako metodo berri hori aurrerapauso garrantzitsua da itzulpen automatikoan, neurona sareek, giza burmuinean inspiratutako algoritmo informatikoen, testu paralelorik gabe itzultzen ikas dezaketela erakusten baitu lehen aldiz. Kasualitatez, beste ikerlan batek, Facebook-eko eta Sorbonako Unibertsitateko (Paris) ikertzaile talde batek egindakoak, antzeko metodo bat proposatzen du. «Harrigarria da —nabarmendu du Artetxek— gure metodoak hain antzekoak izatea. Baina aldiberean pozgarria da, hurbilpen berri hau norabide egokian doala esan nahi baitu». Bi ikerlanak, Science aldizkariak bere berrien edizio digitalean jaso zituenak, arXiv gordailu birtualean argitaratu ziren egun bateko aldearekin, eta apirilean Vancouverren (Kanada) egingo den errepresentazio ikasketaren nazioarteko seigarren biltzarrean (International Conference on Learning Representations - ICLR) aurkeztuko dira.

Bi ikerlanen arteko emaitza konparagarri bakarrean UPV/EHUko IXA taldearen metodoak emaitza zertxobait hobekak lortu zituen, % 15eko BLEU puntuazioa lortuz frantsesa eta ingelesaren artean. Puntuaziorik handiena Google Translate bezalako metodo gainbegiratuak lortzen dute, % 40 inguruko tasarekin, eta pertsona batenak % 50etik gorakoak izan ohi dira. «Hau hasiera baino ez da —argitu du Mikel Artetxek— beraz, ez dakigu noraino iritsiko den ikerketa ildo berri hau».

