



IGOR LETURIA AZKARATE  
Informatikaria eta ikertzailea

# ber2tek

## beste aurrerapauso bat euskararentzako teknologietan

**Poliki-poliki, ordenagailuak-eta ari dira gizakiaren hizkuntzak menderatzen eta, oharkabean, gero eta gehiago komunikatzen gara haiekien lengoia naturala erabiliz eta gero eta gehiago laguntzen digute makinek hizkuntza-arazoekin (itzultzen, zuzentzen...). Baina, euskaraz ere bai? Zorionez, Ber2Tek-en gisako ikerketa-proiektuei esker, euskaraz ere txukun samar egiten dute gailu digitalek.**

**Hizkuntza-** eta **hizketa-teknologiek** ematen diete makinei lengoia naturala ulertu, itzuli edota sortzeko ahalmena. Hizkuntza-teknologiak testuekin aritzeko gaitasunari dagozkio (testuak zuzendu, ulertu, itzuli, kudeatu...); hizketa-teknologiak, berriz, hizketa maneiatzeko gaitasunari (hizketa ulertu, sortu...). Normala denez, hedapen handiena eta, beraz, baliabide gehien dituzten hizkuntzetan daude aurreratu teknologia horiek (ingeleza, gaztelania, txinera...). Euskaraz eta beste hizkuntza gutxiagotu batzuetan, ordea, ez da hain ona egoera. Hala ere, euskaldunok ez gaude kezatzeko moduan: gure egoera minorizatua eta hiztun-kopuru urria kontuan izanik, hain gaizki ere ez dabil euskara, proportzioan behintzat.

### BER2TEK PROIEKTUA

Izan ere, Euskal Herriko hainbat eragile aspalditik ari gara euskararentzako hizkuntza- eta hizketa-teknologiak ikertzen. Esaterako, urte asko daramatzagu elkarlanean proiektu estrategiko batean [Elhuyar Fundazioak](#), [Euskal Herriko Unibertsitateko IXA](#) eta [Aholab](#) ikerketa-taldeek eta [Vicomtech-IK4](#) eta [Tecnalia](#) zentro teknologikoek. Lehen, Eusko Jaurlaritzak Etortek programaren bidez diruz lagundutako hiru proiektutan gauzatu zen elkarlan hori: [Hizking XXI](#) (2002-2004), [AnHitz](#) (2006-2008) eta [BerbaTek](#) (2009-2011) programetan. Elkarlan horren azken fruitua [Ber2Tek](#) proiektua da, zeina 2012tik eta 2014ra garatu baita, Elhuyar I+Gk koordinatuta.

Aipatutako teknologiak ikertzen lan handia egin dugu urte horietan; lehendik garatutako hainbat teknologia hobetzen jarraitu dugu, eta beste teknologia berri asko sortzen. Baliabide orokor asko sortu edo hobetu dira ([corpusak](#), [ontologiak](#), [hiztegiak](#)...); baliabide horiek automatikoki sortzeko teknikak landu dira; euskararen

analisirako tresnak hobetu dira; [itzulpen automatikoan](#) aurrera egin da; [edukiak kudeatzeko teknologiak](#) landu dira; irakaskuntzarako teknologiak garatu dira; [hizketaren sorkuntzan](#) eta [ezagutzan](#) aurrera egin da...

Baina Ber2Tek-en eta haren aurrekoen helburu bakarra ez da ikerketa: teknologia horiek ezagutarazi egin nahi ditugu, eta ikerketaren emaitzak aplikazio bihurtu eta jendearen eskura jarri. Proiektuaren amaiera gisa, teknologia horiek arlo jakin batean egin dezaketek ekarpena erakusteko demo edo erakusle batzuk eraikitzen ditugu. Kasu honetan, teknologik Hizkuntzen Industriaren sektoreari zertan lagun diezaioketen erakutsi nahi izan dugu, hau da, itzulpenaren, edukien eta irakaskuntzaren arloek osatzen duten sektoreari. <http://www.ber2tek.com/eu/demoak> helbidean daude ikusgai demoak.

### APLIKAZIO PRAKTIKOEN DEMOAK

Edukien sektorean egin daitekeenaren erakusgarri, [iritzien erauzketa edo sentimenduen analisiaren](#) teknologia zer den erakusten duen demo bat muntatu dugu. Honetan datza iritzien erauzketa: testu bat hartuta, iritzi subjektiborik baduen eta, baldin badu, iritzi horren polaritatea (positiboa edo negatiboa) zein den erazten du, automatikoki. Aplikazio ugari izan ditzake teknologia horrek; adibidez, enpresek erraz jakiteko euren edo euren produktuei buruz sarean zer esaten den (toki ugaritan eta hizkuntza ezberdinetan). Egin dugun demoan, [Armiarma.eus](#) webguneko [Kritiken hemeroteka](#) hartu dugu, zeinak hainbat komunikabide eta argitalpenetan ateratako euskarazko 5.000 kritika literario baino gehiago biltzen dituen, eta kritika horietako bakoitzari puntuazio bat esleitu zaio, automatikoki, Ber2Tek-en garatu dugun euskarazko iritzien erauzketa teknologia aplikatuta. De-

ber2tek Irakaskuntza Demoa

Ongi etorri, Ber2Tek. Aldatu pasahitza / Itxi saioa  
Gunea kudeatu



Esan duzun zenbakia (771649) honela idazten da: zazpiehun eta hirurogeita hamaika mila, seiehun eta berrogeita bederatzia.

Irakaskuntzarako demoan 3Dko avatar batek euskara irakatsiko digu.

moaren webgunean egile, obra, urte edo bestelako parametroen aukeraketa eginda, grafikoki ikus daitezke puntuazioak, bai eta kritika bera ikusi eta hitz positibo eta negatiboak aztertu ere. Izan ere, hitz horietan oinarritzen da teknologia puntuazioak esleitzeko.

Beste demo batek itzulpenaren arloan egin daitekeena erakusten du, bilatzaile multimedia baten bidez. Gaztelaniaz nahiz euskaraz dauden hainbat bideo hartu dira eta, ahots-ezagutza erabilia, automatikoki transkribatu. Bideo horietako testua eskuratuta, bilaketak egin daitezke haietan eta, nahi izanez gero, instantean jauzi egin bilatzen den hitza esaten den unera. Bideoen transkribapenak gaztelania, euskara edo ingelesera ere itzultzen ditugu, automatikoki, eta azpituduluak erakuts ditzakegu hizkuntza horietan. Itzuli ondoren audioa ere sortzen dugu beste hizkuntza horietan, hizketaren sintesiaren teknologia erabilia; bideo horiek hizlari jakin batzuen hitzaldiak badira, berriz, beste hizkuntzan sortutako hizketa hizlariaren ahotza imitatuz sortzen da, ahotsaren transformazioaren teknologia erabilia.

Azkenik, hizkuntzen irakaskuntzako tutore personal baten demoa egin dugu sektore horretarako. Duela hiru urte, BerbaTek proiektuaren amaieran, antzeko bat egin genuen, baina oraingo honek eginbide eta aukera gehiago ditu; bestealde, mahaigaineko aplikazio bat zen hura, eta

oraingo hau, berriz, on line dago eta edonork probatu dezake. 3Dko avatar bat da demoko tutorea, zeinarekin euskaraz komunikatzen garen, ahoz. Tutoreak gidatzen gaitu automatikoki sortutako aditz-, deklinabide- edo ulermen-ariketan; gure ahoskera ebaluatzen digu; aditz jakin batzuen deklinazioari buruz galde diezaiokegu, bai eta zenbaki jakin bat nola idazten den ere; hiztegian hitz bat bilatzeko esan diezaiokegu; hainbat hiztegitako emaitzak erakutsiko dizkigu...

Muntatu ditugun demo horiek, izenak berak dioenez, demoak besterik ez dira, baina teknologien egungo egoeraren eta egin dezaketenaren gutxi gorabeherako ideia bat izateko balio dute, eta espero dugu benetako tresnetan aplikatuta ikustea laster, lehenago beste teknologia batzuetatik aplikazio errealek sortu diren bezala.

Demo horien bidez ikusiko dugunez, egia da hasieran genioena; euskararentzako hizkuntz- eta ahots-teknologiak nahiko aurreratuta daudela, alegia. Hala ere, bide luzea egin beharko da, oraindik, beste hizkuntza batzuen egoerara iritsi nahi badugu eta egiazki gailu elektronikoe-kin eguneroko bizitzako arlo guztietan euskaraz egin nahi badugu. Ber2Tek proiektua aurrera eraman dugun erakundeok behintzat ez dugu gure lana amaitutzat eman proiektua bukatzean, eta lanean jarraitzen dugu asmo hori noizbait gauzatzeko. ●

“Muntatu ditugun demoek teknologien egungo egoeraren eta egin dezaketenaren ideia bat izateko balio dute”