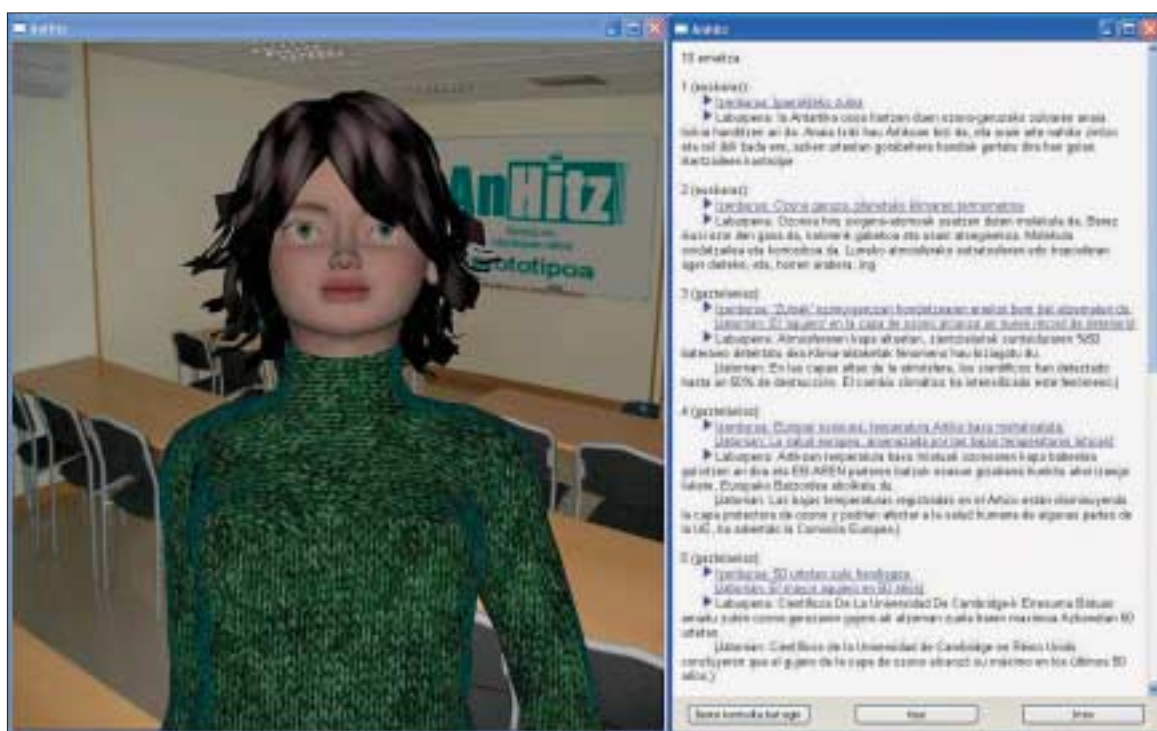


AnHitz, zientzia-aditu birtuala

Kortabitarte Egiguren, Irati

Elhuyar Zientziaren Komunikazioa



Hizketaren teknologiak etengabe ari dira garatzen informatikaren arloan. Interaktibitatea ere puri-purian dago, eta baliteke etorkizunean ordenagailuen mundu birtualekin dugun harremana erabat ezberdina izatea. Euskaraz ere egin ditu urratsak arlo horretan. AnHitz pertsonaia birtuala da horren erakusgarri.

2006TIK 2008RA GARATU DEN ANHITZ PROIEKTUAREN EMAITZA DA azaroaren erdialdean aurkeztu zuten AnHitz demoa edo prototipoa. AnHitz pertsonaia birtual bat da, zientzia-kontuetan aditua, gainera. Euskaraz hitz egiten du. Egiten zaizkion galderak ulertzeko gauza da, eta galderei erantzuteko ere bai. Hor-taz, pertsonaia birtual interaktiboa da, zientzian aditua eta euskalduna.

Erabiltzaileak zientzia eta teknologiari buruzko galderak (Nork asmatu zuen teleskopioa?, Noiz aurkitu zuten penizilina?, Non jaio zen Newton?..) egin diezazkioke sistemari, euskaraz. Era berean, sistemak euskaraz ematen ditu erantzunak. Galderei erantzuteaz gain, hitz bakar edo anitzeko kontsulta bat

egiten bazaio, gai da kontsulta horri dagozkion dokumentuak bilatzeko.

Hitz bakar edo anitzeko kontsultak ere euskaraz egiten zaizkio (ozono-geruza, esnea edateko gaitasuna eboluzioan, energia-iturri berriztagarrien alde onak eta txarrak...), baina kontsulta horiekin zerikusia duten gaztelaniazko eta ingelesezko dokumentuak ere bilatzen ditu, euskarazkoez gain. Gaztelaniazkoak automatikoki itzultzen ditu euskarara sistemak, eta ingelesezkoak jatorrizko hizkuntzan ematen ditu oraingoz, baina laster horiek ere itzuliko ditu euskarara.

Zeregin hori betetzeko, Interneteko zientzia eta teknologiako hainbat gune-

tan egiten du bilaketa. Gainera, nahi izanez gero, testu horiek irakurri egiten ditu sistemak, hiru hizkuntza horietako edozeinetan. Horrez gainera, testuen izenburuan klik eginez gero, haien jatorrizko web gunera sar daiteke erabiltzailea, eta automatikoki itzultakora ere bai, hala nahi badu. Hori guztia lortzeko, jakina, beharrezkoa da teknologia bat baino gehiago uzartzea.

Hamaika teknologia

Aita eta ama bat baino gehiago ditu AnHitz-ek. Izan ere, hainbat enpresak elkarlanean egin duten proiektu baten alaba da. Itxura VICOMTech-ik4 ikerketa-zentroak eman dio, avatarrak edo pertsonaia birtualak sortzeko duen teknologiaz baliatuta. Kasu honetan, emakumezko-itxurako pertsonaia birtual bat garatu du. Aipatzekoa da hiru dimentsiokoa dela, eta emozioak adieraz ditzakeela. Pertsona errearen antza izateak komunikazioa errazten du.

EHUko Aholab taldeak eta Robotiker-Tecnalia ikerketa-zentroak, berriaz, ahotsaren ezagutzan eta sintesian trebatu dute. Horri esker, AnHitz gai da euskalkiak eta hitzak ahoskatzeko hainbat aldaera ezagutzeko, berak euskara batuan hitz egiten badu ere. Galderei erantzuteko sistema eta Interneteko dokumentuak euskarara modu auto-



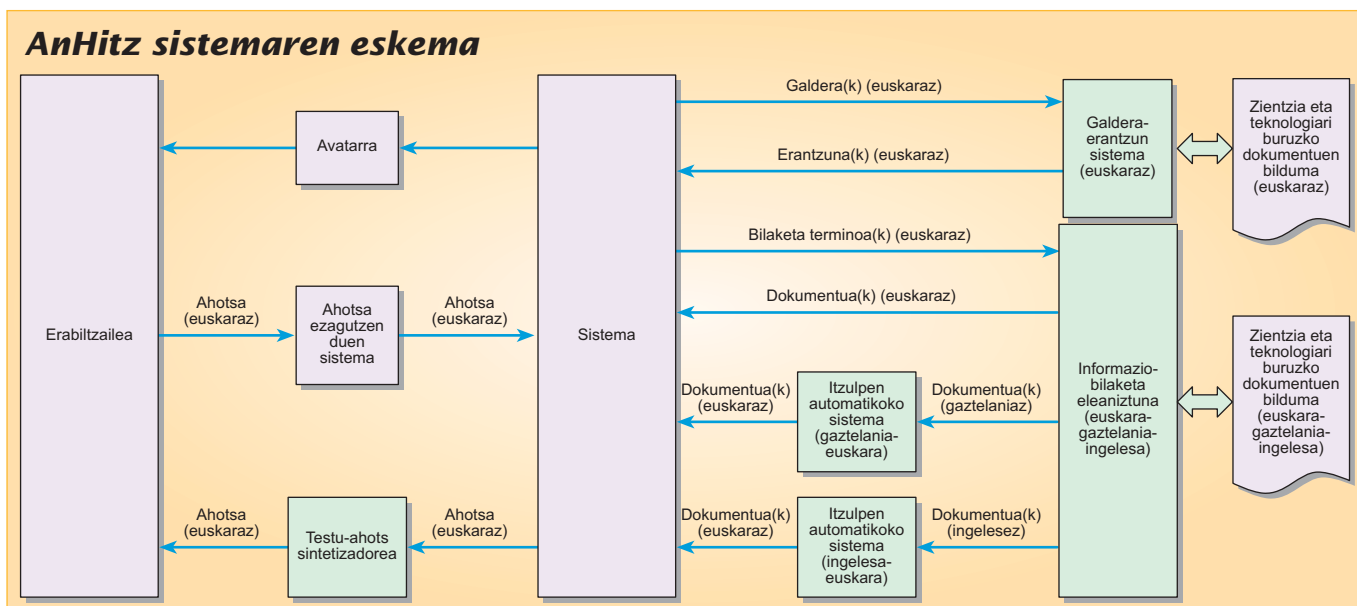
Zientziaren, Teknologiaren eta Berrikuntzaren Astearen, AnHitz ezagutzeko aukera izan zuten Donostiako eta Bilboko zientzia-karpetan izan ziren bisitariek.

“proiektu honetan, lehenengo aldiz, euskararen alorrean dauden hainbat erreminta aplikazio bakarrean bildu dituzte”

matikoan itzultzeko gaitasuna EHUko IXA taldeari zor dizkio. Azkenik, dokumentu eleaniztunak bilatzeko abilezia Elhuyar Fundazioaren I+G taldeak egindako lanaren ondorio da.

Proiektu zabal honetan, erakunde horiek zazpi baliabide, hamahiru tresna eta zazpi aplikazio garatu edo hobetu dituzte euskararen alorrean. Eusko Jaurlaritzaren ETORTEK ikerketa-programaren diru-laguntza jaso du.

Gainera, proiektu honetan, lehenengo aldiz, euskararen alorrean dauden hainbat erreminta aplikazio bakarrean bildu dituzte. Hain zuzen, teknologia hauek uzartu dituzte: *Question Answering* (galderak erantzutea), *Cross Lingual Information Retrieval* (informazio-bilaketa eleaniztuna), *Machine Translation* (itzulpen automatikoa), *Speech Synthesis* (ahotsaren sintesia),



Pertsonak eta gailuak elkarrekin hizketan

Pertsonen eta gailuen arteko elkarrizketa posible izateko, bate-tik, erabiltzaileak hitzez dioena entzun eta deskodifikatzeko (uler-tzeko) gai izan behar du gailuak. Bestetik, emititu nahi den infor-mazioa birkodifikatu (esaldiak sortu) eta erabiltzaileari helaraz-teko (hitzen bidez esateko) gai ere izan behar du.

Hori guztia hitzen bidez egiten denez, hizkuntzek berebiziko garrantzia hartzen dute. Izan ere, erabiltzaileak erabiltzen duen hizkuntza ulertu behar du gailuak, eta horrez gain, buelta-tu beharreko informazioa ere hizkuntza batean edo bes-tean eman behar dio erabiltzaileari gailuak. Horregatik, pertsonen eta gailuen arteko komunikazioan hizkuntzare-kin lotuta sor daitezkeen arazoak konpondu ahal izateko, hizkuntza-teknologiak hartzen dituzte oinarri adituek.

Halaber, informazioaren gizartearen garai honetan, inoiz baino testu, informazio eta jakintza gehiago dago guztion eskura, eta informazio hori guztia egoki tratatzeko tresnak ezinbestekoak dira. Tresna horietan ere beharrezkoak dira hizkuntza-teknologiak.

Gaur egun, hizkuntza-teknologiaren inguruan dauden apli-kazioak bost taldetan sailka daitezke:

- Testuak editatu eta ulertzeko laguntza.
- Testu-kantitate handia erabiltzen den eta kudeatzen duten aplikazioak.

- Itzulpen automatikoa
- Erabiltzaile-gailu interfazeak.
- Hizketaren tratamendua.

Hizkuntza-teknologietan, aurrerapenak beharrezkoak dira era-biltzaileak gailuekin naturaltasun osoz komunikatu ahal izateko. Ikertzaile ugari dihardute lan horretan, eta testuinguru horretan jaio zen AnHitz proiektua.



A. GALARRAGA

Automatic speech Recognition (aho-tsaren ezagutza) eta *Visual Interface* (ikus-interfazeak). Teknologia horien guztien integrazioaren ahalmena ikus-teko, demo bat garatu dute; AnHitz, hain zuzen ere.

“AnHitz proiektuan erabilitako hizkun-tza-teknologia ahalik eta gehienak elkarlanean aritzeko aukera emango digun sistema batean oinarrituta dago demoa” dio Igor Leturia Elhuyar Fun-dazioko I+G taldeko kideak eta demoaren arduradunak. Demo horren bidez, erabiltzaileak eta makinak el-karrizketa bat izan dezakete.

Zientziaren, Teknologiaren eta Berri-kuntzaren Asteko Donostiako eta Bil-boko zientzia-karpetan AnHitz gertutik ezagutzeko aukera izan zuten herrita-rrek. Haur, gazte nahiz helduek egin-dako makina bat galdera eta kontsulta jaso zituen AnHitzek. “Harrera oso ona izan du, eta zenbaitetan jendearengan, haurren batez ere, harridura ere sortu duela esan liteke” gehitu du Leturiak. “Oro har, helduek kontu

*“oraingoz, zientzia-
eta teknologia-
gaietan baino ez da
aditua, eta ez die
galdera-mota
guztiei erantzuten”*

handiz eta hitzak neurtuta egiten ziz-kioten galderak AnHitz, horren guz-tiaren atzean egon litekeen teknologia guztiaz jabetuta, nolabait esateko; haurrek, berriz, beren ezjakintasunean, naturaltasun osoz egiten zizkioten gal-derak, galdera-mota guztiak” gehitu du.

Nolanahi ere, AnHitz prototipo bat da. “Oraingoz, zientzia- eta teknologia-gaietan baino ez da aditua, eta ez die galdera-mota guztiei erantzuten: nor, nor, noiz edo non galdetzaileekin egin-dako galderari erantzuten die ondoen” azpimarratu du Leturiak. “Hala ere,

zalantzarik gabe, hainbat bide berri ire-kitzen dizkigu honek. Izan ere, lite-keena da etorkizunean beste zenbait ezagutza-eremutan ere, turismo-bule-goetan esaterako, aplikatu ahal izatea, bilaketak beste ezagutza-base batzue-tan eginez, noski. Halaber, AnHitz-en atzean dagoen teknologia guztia hobet-zen jarraituko dugu, eta pertsonen eta gailuen arteko komunikazioa errazten jarraituko dugu” dio Leturiak.

Hain zuzen ere, horixe da AnHitz proiektuaren helburua: pertsonak eta gailuak elkarrekintzan aritzeko behar diren euskarazko hizkuntza-teknolo-giak ikertzea eta garatzea, eta ezagu-tzak modu naturalean, intuitiboan eta atseginean kudeatzea. Komunikazio-maila ona eta atsegina lortzeko, hiz-kuntzaren erabilera komunikazioaren alor guztietan sartu behar da: aho-tsean, testuan eta baita multimedia-materialetan ere. AnHitz pertsonaia birtuala horren erakusgarri txiki bat besterik ez da.