

Euskarazko Informazioaren eta Komunikazioaren Teknologiak (IKTak)



Araceli Diaz de
Lezana

Hizkuntza Politikarako Sailburuordetzako Hizkuntza
Plangintzako arduraduna.

"Informazioaren gizartea" kontzeptua entzun genuen, lehenengo aldiz, orain urte batzuk eta, horrekin batera, "Teknologia Berriei" buruz hitz egiten hasi ginen. Internetek, hau da sareen Sareak, bizitza aldatu digu: lan egiteko modua, negozioak egiteko modua, pertonekin erlazionatzeko modua, zerbitzuak jasotzeko modua, informazioa jasotzeko modua... Internetekin batera beste kontzeptu berri bat ere sortu zen "errealitate birtuala", gure ordenagailuaren bidez iristen zaigun beste espazio bat.

Mundu berri horri esker, beharizan berriak sortu dira, neurri handi batean Internetek eraginda, eta, ezinbestez, garapen teknologikoa arlo jakin batean zentratu da bereziki, hizkuntzen garapen teknologikoan hain zuzen ere. Horrela, gaur egun, teknologia horiei "Informazioaren eta Komunikazioaren Teknologiak (IKTak)" deitzen diegu.

Nazioartean, *Information Technology* (IT) izenaz ezagutzen dira. Ikus dezagun nola definitzen duen kontzeptu hori *Information Technology Association of America*-k: "*the study, design, development, implementation, support or management of computer-based information systems, particularly software applications and computer hardware...*". *Information Technology is a general term that describes any technology that helps to produce, manipulate, store, communicate, and/or disseminate information.*

[Iturria: Wikipedia] (Hardwarea eta softwarea baliatuz, ordenagailu bidezko informazio-sistemak aztertzea, diseinatzea, garatzea, inplementatzea, mantentzea eta kudeatzea... Informazioaren Teknologiak termino orokorra da eta informazioa sortzen, manipulatu, biltegitzen, komunikatzen edota hedatzen laguntzen duten teknologiak deskribatzen ditu).

Beraz, Informazioaren eta komunikazioaren aroan bizi gara eta, hori dela eta, unibertsitateak, ikerketa-zentroak eta enpresak hizkuntzen trataera berezia egiten duten teknologiak garatzen ari dira, "infoingeniaritza linguistikoa" edo "hizkuntzen teknologiak (*Language Technologies*)", hain zuzen. Negozio bide eta merkatu berriak agertu dira, eta, horiek bultzatuta, aipatu teknologiak aurrerapauso handiak egiten ari dira. Batik bat honako hauek dira:

- Ahotsa edo hizketa ezagutzea (**Speech-To-Text** eta **Automatic Speech Recognition**)
- Ahotsaren edo hizketaren sintesia (**Text-To-Speech**)
- Karaktereen ezagupen optikoa (**Optical Character Recognition**)
- Itzulpen automatikoa (**Text-To-Text** eta **Speech-To-Speech**)
- Ahozko datuen konprensioa
- "Data / audio mining" (dokumentazio-multzo handiak kudeatzea)

Gaur eguneko komunikazioan teknologia lehentasuna da, herri baten garapen-maila adierazten baitu. Teknologiarri esker erraz komunikatzen gara, geografiaren eta denboraren oztopoak desagertzen direlako. Dena den, teknologiaren eragina beti erabiltzailearen eskuetan izango da, hau da ematen dion erabileraren arabera izango da: informazioa lortu, entretenimendua bilatu, lan egin, ikasi, prestakuntza jaso... edo, bestelako helburuetarako erabili: iruzurra egin, pertsonak kontrolatu eta kaltetu, datuak manipulatu eta faltsutu ...

Edozein kasutan, IKTen garapena erabiltzaileek markatuko dute, beraiek erabakiko baitute zelako produktuak eta edukiak nahi dituzten une orotan. Horrela, esaten da teknologiek eragin handia dutela gizartea eraikitzeke prozesuan. Horregatik, oso garrantzitsua da audientzia formatzea, izan ere, erabiltzaileek sormena zenbat eta gehiago erabili, orduan eta hobekiago izango dira garatuko diren produktuak eta gutxiago merkatu-bazterketak.

Informazioaren eta Komunikazioen Teknologien garrantziaz jabetuta, gobernuek eta administrazioek horiek garatzeari begirako politikak ari dira gauzatzen. Izan ere, badakite, batetik, garapenaren ondorioz gizarte-, ekonomia- eta politika-aldaketak gertatzen ari direla, eta, bestetik, teknologiak erabiltzeak ahalmen ia mugagabea eskaintzen duela informazioa eskuratu, bildu, prozesatu eta partekatzeke.

1. Euskarazko Informazioaren eta Komunikazioaren Teknologiak.

"Euskadi 2000 Hiru Ekimena"ren esparruan Eusko Jaurlaritzak hainbat ekimen sustatu zituen euskal gizartea bete-betean sar zedin Informazioaren Gizartean; horretarako, espazio digitala pertsonen zerbitzura jarri zuen.

Horretaz gain, hainbat plan bideratu du Eusko Jaurlaritzak Euskadin Informazioaren eta Komunikazioaren Teknologiak garatzeko eta euskara berauetan integratzeko:

1999, Euskara Biziberritzeko Plan Nagusia.

Planak hurrengo hamar urteotarako honako helburu hauek ezarri zituen: euskararen ingeniartzaren industria bultzatu; merkatuan zabalkunde handia duten produktuen lokalizazioari lehentasuna eman; erabiltzaile-komunitate prestatua eta aurreratua sortu; edizio- eta itzulpen-plataforma edo sareak sortu, baita denen eskura egon daitezkeen hiztegi elektronikoak ere; euskara ikasteko multimedia-sistemak garatu; eta hizkuntza-analisirako oinarritzko tresnen azpiegitura sortu.

2001-2004, Zientzia, Teknologia eta Berrikuntzarako Plana.

Hizkuntza-industrien programaren barruan, Planak lehentasunezko sei esparru hauek bultzatu nahi zituen:

- Hizkuntza-ingeniartzaren oinarritzko tekniken garapena.
- Baliabide linguistikoaren garapena.
- Bitarteko aurreratuen erabilera bultzatzea eta plataforma gaitasun-emaileen esperimentazioa.
- Informazioa bildu, berreskuratu eta bilatzea.
- Ahotsaren ezagutza automatikoa, ahotsaren konpresio digitala eta testua ahots bihurtzea.
- Testuaren tratamendurako funtzionalitate lagungarriak.

- Testua bulegotika edo ofimatikaren testuinguruan tratatzea, bai eta testutik testurako aldaketan ere (itzulpena).

2002-2005, Euskadi Informazio Gizartean Plana (PESI).

"Euskadi Sarean" atalean euskararen aldeko ekimenak aipatzen dira:

"Hizkuntza-baliabideen garapena sarean, euskarazko edukien garapena eta sustapena sarean, teknologia berrien alorreko proiektu berritzaileen sustapena eta euskarari aplikatutako ahotsaren, itzulpen automatikoaren eta testu-tratamenduaren teknologiak".

2003-2005, Informatika eta Telekomunikazio Plana.

Aurreko planetan aipatutako ekimenetatik sortutako informazio-sistemen proiektuak eta beste batzuk hartu zituen bere baitan plan honek, honako jardun-ildo hauek aintzat hartuta:

- Barruko kudeaketa hobetzea
- Administrazio elektronikoa
- Azpiegiturak eta antolakuntza

PCTI 2010, Zientzia, Teknologia eta Berrikuntzarako Plana

8.1.7 eta 8.1.8 ataletan Informazioaren teknologien eta Hizkuntzen industriren gaineko helburuak finkatzen dira.

2008, Euskadi Informazioaren Gizartean Plana 2010: Euskadiko Agenda Digitala

"Euskadi Sarean" atalean honako hau zehazten da:

Euskadi sarean egon dadin ziurtatzea, ingurune digitaletan euskal kultura eta hizkuntza babestuz eta Interneten erakuslehoan euskal turismo-baliabideen eskaintza aurkeztuz.

Programak honako hauek dira:

- Euskal kulturaren eta euskal kultura-ondarearen aldeko zerbitzuak
- Euskaraz erabateko esperientzia digitala ahalbidetzeko zerbitzuak
- Euskadik kalitatezko helmuga turistiko gisa duen irudia indartzeko zerbitzuak.

Plan horiekin batera, aipatu behar dira Kultura Saileko Hizkuntza Politikarako Sailburuordetzak IKTak sustatzeko urtero kaleratzen duen diru-laguntzen deialdia, bereziki euskarazko IKTen arlorako produktuak finantzatzen dituen, eta Industria Sailak, ikerketa sustatzeari begira, abian dituen programak: ETORTEK, SAIOTEK eta INTEK. Programa horiek Zientzia, Teknologia eta Berrikuntza Sareko agenteei daude zuzenduta. Foru-aldundiek ere diru-laguntzak ematen dituzte teknologia horiek sustatzeko.

Hori horrela, IKTen arloan, azken urteotan, administrazioen ekimenez, euskara integratzen duten proiektu asko garatu da; horri gehitu behar zaio agente teknologikoek arlo horretan egindako eta egiten ari diren lan handia.

Labur-labur eta oro har aipatuko ditugu euskaraz garatu diren produktuak eta baliabideak:

- **www.euskadi.net/euskara_soft** helbidean hainbat produktu eskura daiteke doan: Windows Vista eta XP (interfaze-paketea) eta Linux sistema eragileak; MS Office 2007 eta 2003 (interfaze-paketea) eta OpenOffice.org bulegotika-paketeak; plugin-ak (Elhuyar hiztegiak, UZEIren sinonimoen hiztegia,

QuarkPress-erako hitz-zatitzailea; OCR 1.1); web garapenerako Miatu erraz liburutegia; Xuxen zuzentzaile ortografikoaren bertsio guztiak (kode irekikoa ere bai); eta enpresentzako kudeaketako softwareak (SAP 4.6c bertsio bost modulu; Sugar CRM 4.2.0).

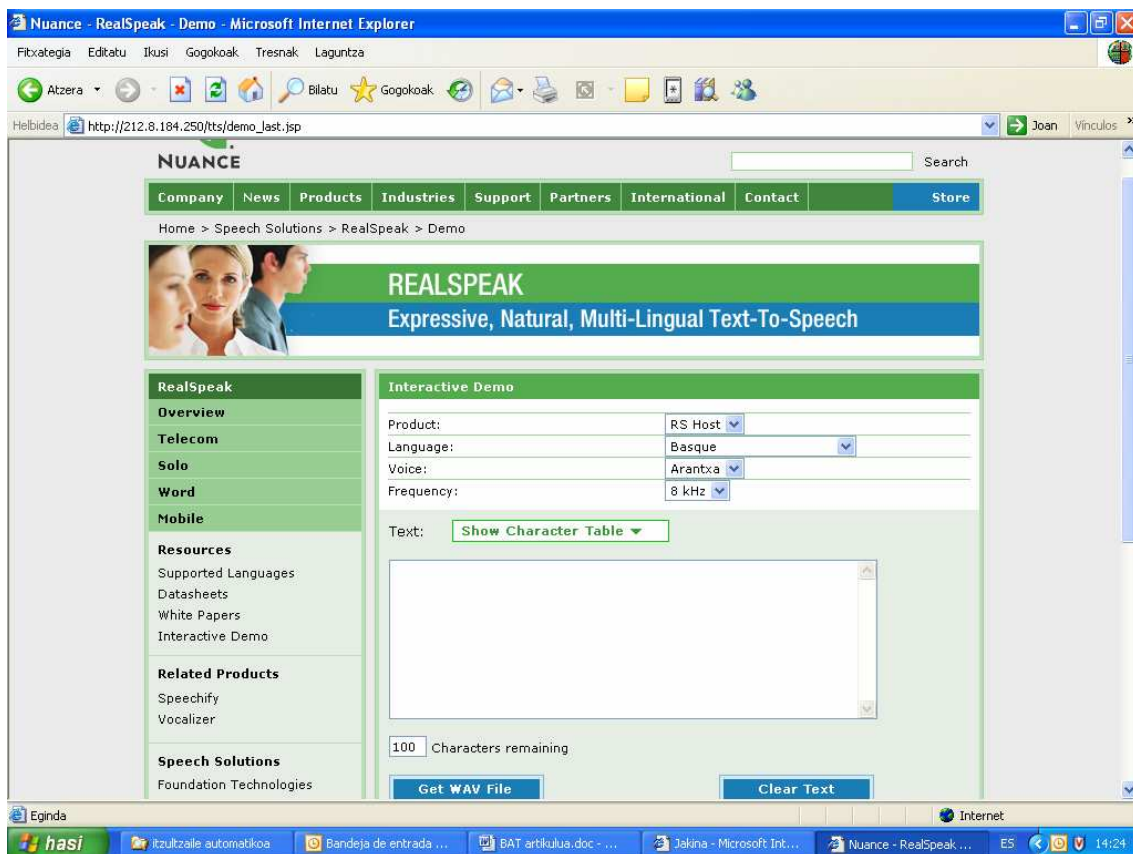
Euskarazko softwarea deskargatzea

Gune honetan ordenagailuarekin euskaraz lan egiteko hainbat oinarriko tresna eskaintzen dira doan

- Xuxen zuzentzaile ortografikoa**
Xuxen III EDBLko (Euskararen Datu Base Lexikala) datuekin; Xuxen II
- Bulegotika**
MS Office; OpenOffice.org
- Sistema eragileak**
MS Windows XP-aren interfaze-paketea; EusLinuxDebian 2005; Mandrake Linux 10
- Kudeaketako softwarea**
SAP 4.6c; Sugar CRM 4.2.0
- Plugin-ak**
QuarkPress-erako hitz-zatitzailea; Elhuyar hiztegiak; UZEIren sinonimoen hiztegia; OCR 1.1
- Web Garapenerako Liburutegia**
Miatu Erraz bilaketa liburutegia

Azken eguneratzea: 2007/11/26

- Ahotsa edo hizketaren sintesia eta ezagutza: Eusko Jaurlaritzako Kultura Sailak hitzarmena izenpetu zuen Nuance enpresarekin (lehen Scansoft) euskararako motoreak garatzeko **www.nuance.com**.



- Hizkuntza-baliabideak: hiztegi elektronikoak (Hiztegi Batua, Elhuyar, 3000, Morris, Euskal Hiztegia, Sinonimoak, Harluxet hiztegi entziklopedikoa...); datu-base lexikalak (EDBL); corpusak (corpus estatistikoa, Zientzia eta Teknologia corpora, literatura-corpusak); ahotsaren teknologietan integratzeko baliabideak (datu-base akustikoak, lexiko fonetikoak, testu-corpusa); itzulpen-memoriak; aldizkari ofizialen corpus etiketatua; terminologia banku publikoa (Euskalterm); toponimia (EAEko datu-basea) eta onomastika (Euskaltzaindiaren izendegiak)...
- Hizkuntza-tresnak: bilatzaileak, lematizatzaileak, analizatzaileak, zuzentzailea, etiketatzaileak, thesaurusak, erauzketa automatikoak egiteko tresnak, itzulpen automatikorako tresnak...
- e-prestakuntza: online ikastaroak (masterrak, ikastaroak...).
- Foroak, eztabaidaguneak, buletin elektronikoak.
- Eduki digitalak (prentsa, aldizkariak...)
- ...

Gainera, 2008ko uztailean, Hizkuntza Politikarako Sailburuordetzak lehiaketa publikoaren bidez, testu orokorrak gaztelaniatik euskarara itzultzeko erreminta bat garatzeko kontratua esleitu du. Lanak hamabost hilabetetan zehar luzatuko dira

kontratua adjudikatu zenetik, eta sortuko den tresna honako esparru hauetan erabili ahal izango da:

- www.euskara.euskadi.net gunetik testu laburrak eta web guneak itzultzeko.
- Eusko Jaurlaritzako intranetaren bidez formatu desberdinetako dokumentuak itzultzeko (sistema hobetzeko erabiliko direnak).

Hori lehenengo urratsa izango da, geroago euskaratik gaztelaniara itzultzeko tresna garatu beharko delako. Eta horren ostean euskara-ingelesa eta euskara-frantsesa. Euskaraz orain arte egindako saioak gaztelaniatik euskarara egin dira, kontuan izanda abiapuntu-hizkuntza garatuta egotea garrantzitsua dela.

2. Euskarazko Informazioaren eta Komunikazioaren Teknologien (IKTen) Inbentarioa.

Hizkuntza Politikarako Sailburuordetzak (HPS) *Hizkuntza politika aurrera begira* argitalpenean 2005-2009 epealdirako egitasmoa definitu zuen (Eusko Jaurlaritza, Vitoria-Gasteiz, 2005, ISBN 84-457-2407-X).



Bertan aipatzen diren neurrien artean, 23.ean honako hau esaten da:

Informazioaren eta Komunikazioaren teknologia berrietan euskarazko baliabideak, tresnak eta aplikazio berriak garatzea. Halaber, euskararen presentzia areagotzea Interneten, eduki digitalak sortuz... Behar-beharrezkoa irizten diogu, nolana ere, alor honetan euskarak gaur egun duen presentziari eta teknologia berri hauetan jadanik badiren euskarazko baliabideei buruzko azterketa eta erradiografia bat egiteari, eta etorkizunean urratu beharreko ildo zehazteari.

Hori horrela, HPSk **IKTen Inbentarioa** osatzeari ekin zion 2006an. Ekimen horren bidez, honako helburu hauek lortu nahi zituen HPSk:

- Egindako edo egiteko bidean diren ekimenen diagnostikoa eta mapa izatea.

- Orain arte egindakoaren indar-guneak eta ahuleziak antzematea eta aurrera begirako lehentasunak definitzea.
- Eusko Jaurlaritzaren plan estrategikoekin bat letozkeen proiektuak eta ekimenak identifikatzea, betiere balio erantsia eta eraginkortasuna kontuan.
- Azken finean, teknologia berrietan euskararen presentzia areagotzeko informazio kualifikatua eskuratzea eta ezagutaraztea Euskal Adierazleen Sistemaren bitartez:
 - Euskararen presentzia teknologia berrietan
 - Teknologia Berriek euskarari eransten diotena

Eman beharreko lehenengo urratsa IKTen arloa sailkatzea izan zen. Izan ere, proiektuak Inbentarioan sartzeko behar-beharrezkoa zen azpieroju jakin baten barruan kokatzea:

**Informazioaren eta Komunikazioaren Teknologien sailkapena
(IKT Inbentarioa)**

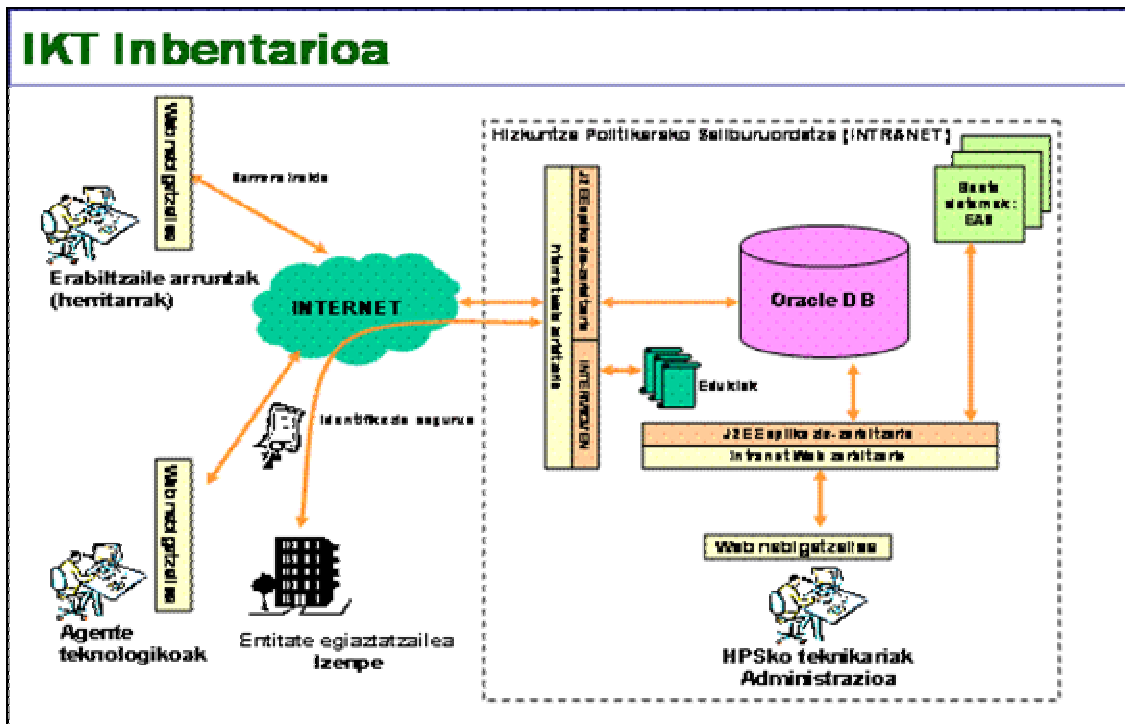
www.euskara.euskadi.net/ikt_inbentarioa

<p>HIZKUNTZA-BALIABIDEAK ETA TRESNAK</p> <ul style="list-style-type: none"> - ahots-baliabideak - lexikoak - hiztegiak - zuzentzaileak - lematizatzaileak - dokumentu eleanitzen edizioa - OCR sistemak - banku terminologikoak - ontologiak - analizatzaileak - etiketatzaileak - tesaurusak - datu-base lexikalak - corpusak - bestelako lanabesak 	<p>INFORMAZIOAREN ETA JAKINTZAREN KUDEAKETA</p> <ul style="list-style-type: none"> - dokumentazioa berreskuratzea - informazioa eraztea - laburpen automatikoak egitea - dokumentazioa sailkatzea - dokumentazioa berbideratzea - dokumentazioa pilatzea - dokumentazioa iragaztea - testuak automatikoki sortzea - bestelakoak
<p>PRESTAKUNTZA / E-LEARNING</p> <p>ITZULPENGINTZARI LAGUNTZEKO TRESNAK</p> <ul style="list-style-type: none"> - itzulpen-memoriak - Itzulpen automatikorako tresnak - bestelakoak 	<p>EUSKARAZKO SOFTWAREA</p> <ul style="list-style-type: none"> - bulegotika edo ofimatika - enpresa-kudeaketa - erabilera orokorreko tresnak - irakaskuntza - hizkuntza - aisialdia - Internet lanabesak - softwareari loturiko informazio teknika - bestelakoak
<p>SISTEMA INTERAKTIBO AURRERATUAK</p> <ul style="list-style-type: none"> - ahots-ezagutza automatikoa - galderak erantzuteko sistemak - ahotsaren bidezko identifikazio biometrikoa - ahotsaren sintesia 	<p>EDUKIEN INDUSTRIA (euskarazko eduki digitalak sortzea)</p> <p>E-INKLUSIOA (ezintasunak dituztenentzako aplikazioak)</p> <p>BESTERIK</p>

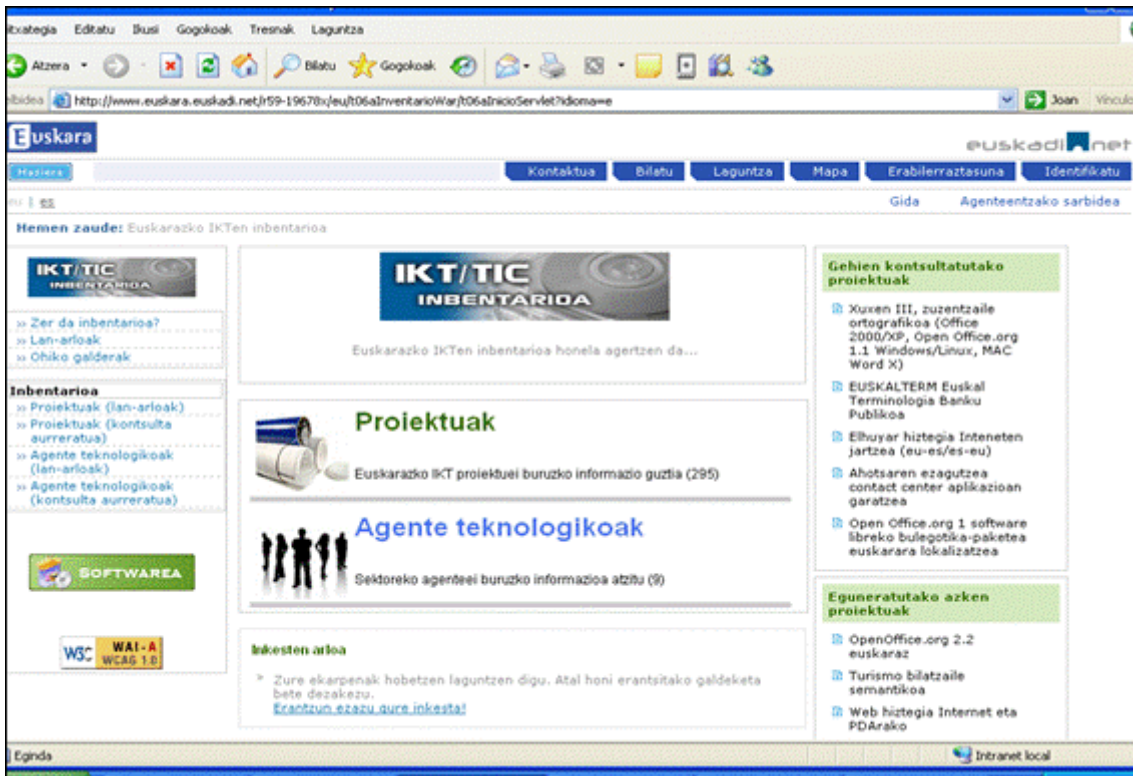
Hurrengo urratsa EAEko agente teknologikoekin harremanetan jartzea izan zen; izan ere, euskarazko IKTen inguruko informazioa biltzeko, erabat beharrezkoa eta

ezinbestekoa zen beren ekarpena. Horrela, hasiera batean 15 agente emandako informazioarekin jarri zen martxan euskarazko IKTen Inbentarioa: Deustuko Unibertsitateko DELi Taldea, Elhuyar Fundazioa, Euskal GNU Elkarte, Eleka Ingeniaritza Linguistikoa, S.L., Euskal Herriko Unibertsitateko IXA Taldea, Eusko Ikaskuntzaren Asmoz Fundazioa, Eusko Jaurlaritzako HPS eta IVAP, Bizkaiko eta Gipuzkoako Foru Aldundiak, Grupo Voice, Plazagunea SL, Robotiker, UZEI eta Vicomtech.

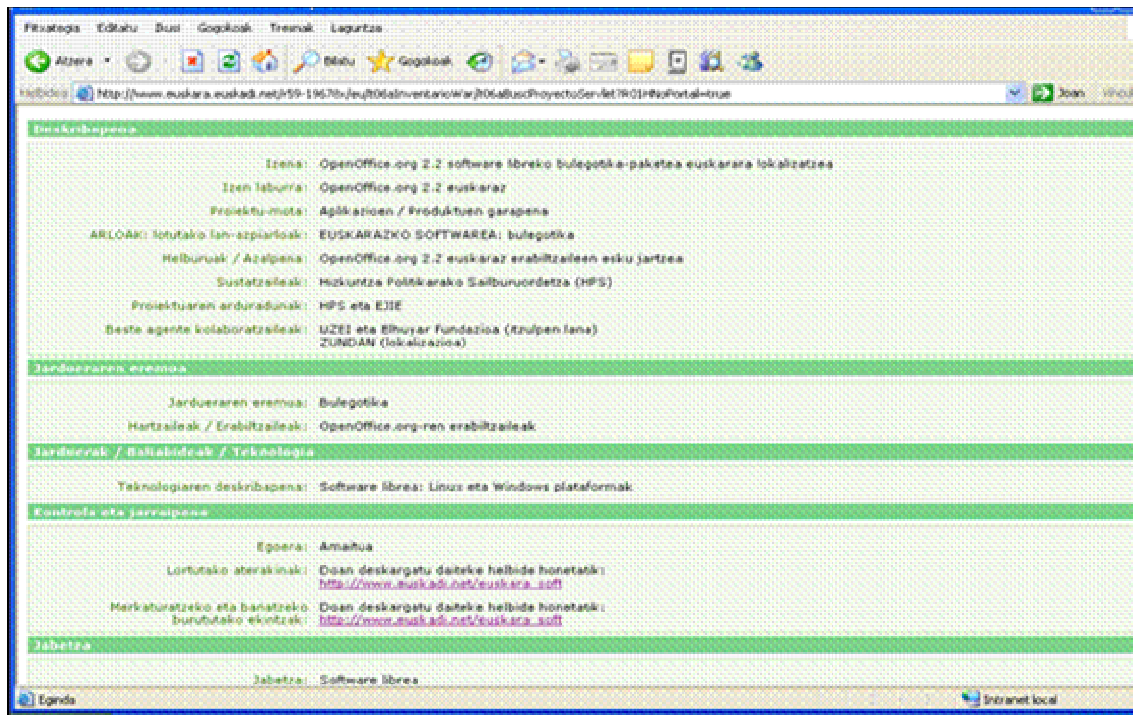
Baina Inbentarioa eguneratuta eta bizirik mantendu behar denez, HPSk aplikazio bat garatu zuen Interneten bidez erabiltzaile guztien eskura jartzeko: alde batetik, kontsulta orokorra eskaintzen da edozein proiektu edo produkturen gainean informazioa bilatzeko eta, bestetik, agente teknologikoen zuzeneko sarbidea dute beren proiektuak eguneratzeko edota proiektu berriak sartzeko. Inbentarioan agente teknologiko bezala sartzeko IZENPE txartela behar da (herritarrarena) eta, gainera, sisteman alta eman behar da. Argibide horiek aplikazioan bertan azaltzen dira. Interneteko aplikazioaz gain, intraneteko dago HPSk Inbentarioa kudeatzeko, besteak beste, iristen den informazioa baliozkotzat emateko. Horrela ziurtatzen da informazioa homogenea eta zuzena dela. Ikus dezagun IKT Inbentarioaren egitura:



Informazioaren eta Komunikazioaren Teknologien Inbentarioa www.euskara.euskadi.net/ikt_inbentarioa helbidean dago, aipatu bezala, erabiltzaile arruntek kontsultatzeko eta agente teknologikoen beren proiektuen gaineko informazioa txertatzeko:

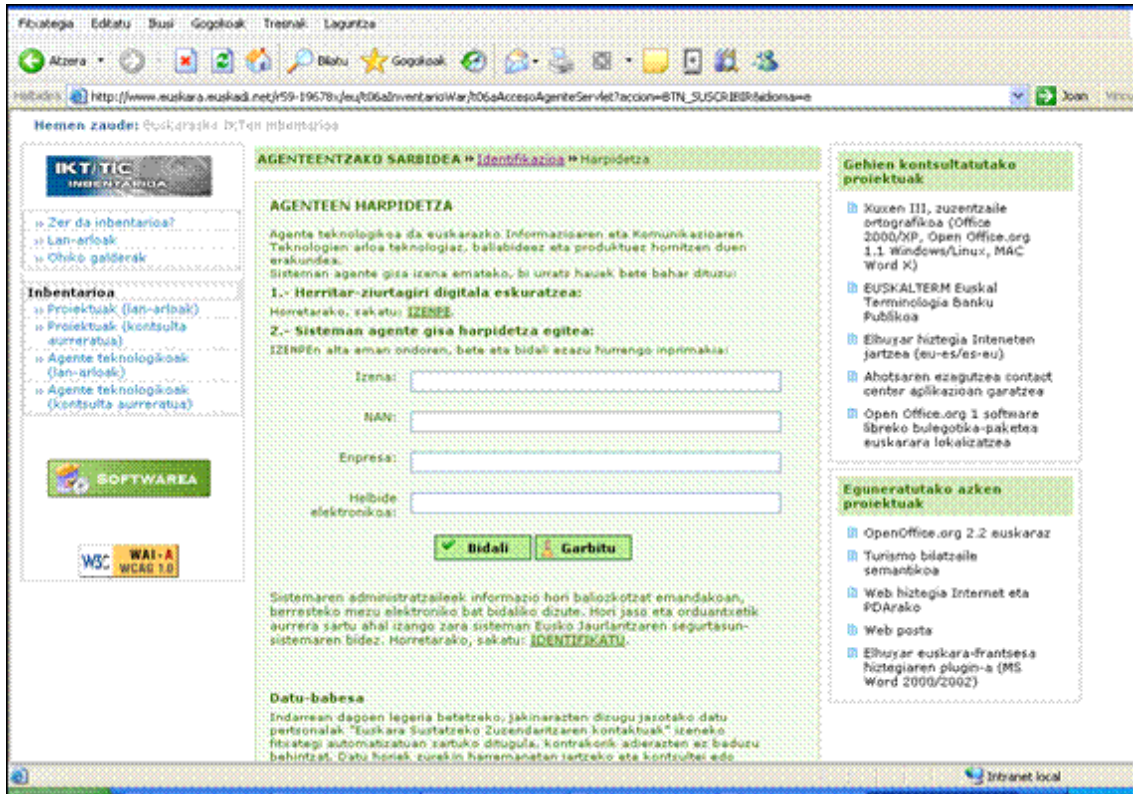


Proiektu jakin bati buruzko kontsulta eginez gero (OpenOffice.org, kasurako), honako informazio hau lortuko du erabiltzaileak:



IKT Inbentarioa bizirik mantenduko bada, zalantzarik gabe, agente teknologikoen ekarpenei esker izango da. Euskarazko tresnen kontsumitzaileek jakin behar dute

momentu orotan zer dagoen garatuta euskaraz, zeinek egin duen eta non lor dezaketen horri buruzko informazioa. Azken finean, helburua da euskarazko IKTen merkaturia sustatzea eta areagotzea. Arestian esan bezala, agente teknologiko gisa Inbentarioa erabiltzeko behar-beharrezkoa da identifikazio segurua (herritarraren IZENPE txartela) eta sisteman alta ematea:



Agente teknologikoen proiektu bakoitzari buruz eman behar duten informazio osoa ez da Interneten bistaritzen, eremu batzuk barruko kudeaketarako bakarrik erabiltzen baitira.

Beraz, momentu honetan hauxe da IKT Inbentarioan jasota dauden proiektuen/produktuen kopurua:

- Hizkuntza-baliabideak eta tresnak: 70
- Prestakuntza / e-learning: 7
- Itzulpengintzari laguntzeko tresnak: 15
- Sistema interaktibo aurreratutako: 19
- Informazioaren eta jakintzaren kudeaketa: 24
- Euskarazko softwarea: 198
- Edukien industria (euskarazko eduki digitalak sortzea): 9
- E-inklusioa: 5
- Besterik: 4

Proiektuak/Produktuak eremu batean baino gehiagotan sailkatu daitezkeenez, errepikapenak agertzen dira.

3. Euskararen Aholku Batzordearen (EAB) IKT Batzordea.

Urriaren 16ko Euskararen Aholku Batzordeari buruzko 176/2007 DEKRETUAK (EHAA, 204 zk.) honako hau ezartzen du: "Aholku Batzordearen Batzorde Osoak edo Kultura sailburuak batzorde-atal bereziak sortzea erabaki dezakete, gai zehatzei buruzko lanak eta txostenak egiteko. Batzorde Osoaren Erabakia edo Kultura sailburuaren Agindua beharko da horretarako".

Hori horrela, Kultura sailburuaren 2007ko urriaren 30eko AGINDUAK (EHAA, 235 zk.) Euskararen Aholku Batzordeko batzorde-atal bereziak sortu zituen. Besteak beste, "Informazio eta komunikazio teknologiak" izena duen batzorde-atala sortu zuen. Hauek dira Aginduak ezartzen dizkion egitekoak:

"Hizkuntza Politikarako Sailburuordetzaren 2005-2009 aldirako egitasmoa jasotzen duen Hizkuntza Politika aurrera dokumentuko 23. jardunbidearen garapenerako egoki diren iradokizun eta proposamenak egitea da, besteak beste, batzorde-atal honen xedea. Izan ere, informazioaren eta komunikazioaren teknologia berrietan euskarazko baliabideak, tresnak eta aplikazio berriak garatzea ezinbestekoa da, bai eta euskararen presentzia areagotzea Interneten, eduki digitalak sortuz. Hori hala, behar-beharrezkoa da ere alor honetan euskarak gaur egun duen presentziari eta teknologia berri hauetan jadanik badiren euskarazko baliabideei buruzko azterketa eta erradiografia bat egitea, eta etorkizunean urratu beharreko ildoak eta estrategiak proposatzea".

EABko IKT Batzordea martxan da eta 2008ko Jarduketa-plana onartua du. Batzordearen lanaren abiapuntua IKTen Inbentarioa da, Inbentarioko informazioa oso baliotsua baita euskarazko IKTen hutsuneak non egon daitezkeen jakiteko eta, hortaz, zeintzuk diren arlo horretako lehentasunak. Hauek dira Jarduketa Planean jasotakoak:

- INTERNETen inguruko azterketa egin, lanak proposatzeko: euskararen erabilera ebaluatu (zenbat web gune euskaraz, euskal blogosfera, zenbat jendek nabigatzen duen euskaraz...), Interneteko euskarazko tresnak (garatzen dituzten taldeak, produktuak...), euskarazko edukiak sarean, gazteentzako ekimenak...
- Hizkuntza-baliabideak eta tresnak: bi corpus eratzeko proiektuak definitu: XXI. mendeko euskararen erreferentzia-corpusa, eta itzulpen automatikoan erabiltzeko corpus elebidun etiketatua.
- Honako arlo hauetan premiak aztertu eta proiektuak proposatu: sistema interaktibo aurreratuak (ahots-teknologiak, galderak erantzuteko sistemak...); euskarazko softwarea (enpresa-mundurako aplikazioak euskarara lokalizatu, software librea nahiz jабeduna); Informazioaren eta jakintzaren kudeaketa; eta e-inklusioa (ezintasunak dituztenentzako aplikazioak).

4. Etorkizunari begira zer?

Hizkuntza handien kasuan, bereziki bi arrazoiengatik ari dira Informazioaren eta Komunikazioaren Teknologiak garatzen eta bultzatzen:

- Ekonomikoak: helburua da teknologian inbertitzea, gero kostua murrizteko. Teknologia horien erabiltzaileek badakite, oraingoz, ez direla perfektuak eta hainbat kasutan (oso desberdinak diren hizkuntzen arteko itzulpena, esaterako) emaitzak ez direla onak. Kalitatearen kontzeptua, ordea, aldatzen ari da eta orain

beste bi kontzeptuekin batera ulertu behar da: "kantitatea, kalitatea eta prezioa" (*quantity, quality, price*). Irizpide hori erabiltzen du bete-betean Europar Batasunak, bertan hizkuntza ofizialen kopurua handitzeaz batera, kostuak izugarri areagotu direlako.

- Hizkuntza-aniztasuna (*multilingualism*): informazioa eta komunikazioa gero eta konplexuagoak dira. Hizkuntzen teknologietan hizkuntza bakarra edo bi hizkuntza bakarrik erabiltzeko praktikak erabat baztertuta geldituko dira.

Eta, aurrera begira, zeintzuk izango dira ikerketa- eta negozio-arloak IKTei dagokienez? Enpresak eta ikerketa-zentroak honako arlo hauen inguruan ari dira fokalizatzen beren aurreikuspenak:

Ahotsaren teknologiak.

- Eskuko telefonoak: datozen urteotan horietan ikusten dute enpresek negozioerik inportanteena (*key technology area*). Aurreikuspenen arabera, hurrengo urteotan izugarri handituko da eskuko telefonoen kopurua. Beraz, sistemak garatu behar dira hizkuntza aniztan zerbitzuak telefonoz emateko (negozioei buruzko informazioa, hirien egoerari buruzko datuak, helbideak, erosketak, erreserbak, eguraldia, telebistako programak, kirolak, jokoak...). Sistema horiek eragina izango dute munduko biztanleriaren frakzio handi batean. Telefonoak Internetekin konparatuz gero, aplikazioak garatzeko orduan, hainbat informazio ematen du: pertsonaren generoa (emakumezkoa/gizonezkoa), adina, heziketa-maila, arraza... Esaterako, esaten da emaitzak oso desberdinak direla gizonezkoekin edo emakumezkoekin erabiliz gero. Beraz, sistema horietan giltzarria izango da interfaze egokiak, multimodalak, egitea. Gainera, telefono bidezko zerbitzu horiek, itsuentzat esaterako, oso erabilgarriak izango dira.
- Hitz egiteaz batera hizketa ulertzea (*spoken language understanding*), eta, gainera, komunikazio-zirkuitu eleanitz batean.
- Ahotsa edo hizketa ezagutzea: ahotsa kontrolatzea, emozioak gehitzea, sistema interaktiboak garatzea...
- Hizketatik hizketara itzulpena (*Speech-to-Speech Translation*): ahoko teknologia eta itzulpen automatikoaren teknologia erabilia.

Itzulpen automatikoa (*Machine Translation MT*).

Praktikak esaten du oso antzekoak diren hizkuntzen arteko itzulpenaren emaitzak oso onak direla (%95); oso desberdinak diren hizkuntzen artean, berriz, emaitzak, oraingoz, ez dira onak (%65 gehienez ere). Dena den, arestian esan bezala, kalitatearen kontzeptua aldatzen ari da eta, orain, beste parametro batzuekin batera neurtu behar da, hau da, kantitatearekin eta prezioarekin. Itzulpen automatikoaren kontsumitzaileek ez dute perfekzioa bilatzen; bereziki, ulertzen ez duten testu baten inguruko gutxi gorabeherako informazioa nahi izaten dute lortu. Beraz, esango genuke nahiago dutela informazio inperfektua izatea, ezertxo ere ez izatea baino. Hori, jakina, informazio orokorrari buruz hitz egiten badugu. Argi dago, bestelako testuinguruetan, hau da testu ofizialen edota antzeko testuen kasuetan, kontu handiz erabili behar direla sistema horiek eta, azken bezeroari eman aurretik giza zuzenketa beharrezkoa dela. Beraz, planteamendua da, erabiltzeaz batera, modu kontrolatu batean, sistema hobetzen joatea, hots, kalitatearen kontrola egin behar da. Erantzuteko denbora ere neurtu behar da. Egia da Internetekin "itzulpen azkar eta merkearen eskaera" handitzen ari dela.

Azken honetan, ebaluatze-sistemak (*BLEU* esaterako) erabiltzen ari dira itzulpen automatikoaren sistemak ebaluatzeko.

Informazioa berreskuratzea (*Information retrieval*).

Informazio garrantzitsua modu azkarrean eta eraginkorrean bilatzeko tresnak dira. Horretarako, hizkuntzak gurutzatu egiten dira, hau da bilaketa eleanitzak egin daitezke (*cross-lingual retrieval*). Hor kokatzen dira liburutegi digitalak ere, hau da hizkuntza desberdinak erabilia, munduko dokumentu inportanteak edo bestelako altxorak erabiltzaileen eskura jartzeko sistemak. Teknologia honek galdera-erantzuna sistema erabiltzen du bai eta ontologiak eta thesaurusak ere. Informazioa erlazionatu, sailkatu eta laburtu egiten du. Azken finean, web semantikoaren teknologia da: kontzeptuak, entitateak eta gertaerak izaten ditu kontuan, egiturazko azterketak egiten ditu eta etiketa semantikoak jartzen ditu.

Internet

Orain gutxi pertsona batek esan zuen Internet dela "agora" berria. Hau da, Informazioaren eta Komunikazioaren aroak Interneten du ardatza. Azken urteotan Sarearen bilakaera azkarra ikusi dugu. Azkeneko honetan, web sozialetik (2.0) web semantikora (3.0) pasa gara. Gero eta eduki gehiago dago Interneten eta hizkuntza askotan. Beraz, eduki konplexu horiek kudeatzeko sistemak behar dira. Eta sistema horiek hizkuntza, makina bidez ikastea eta jakintza irudikatzen teknologia behar dute testuinguru berrietan aplikatzeko. Semantikan oinarritzen dira, hau da, kontzeptuetan, kontzeptuen arteko erlazioetan, eta ondorioz ontologietan, thesaurusetan. Web semantikoan ere galdera-erantzuna sistema erabiltzen da. Esan behar da, Interneten eduki eta zerbitzuaren bereizketa desagertzen ari dela.

Baina **hizkuntzen teknologien** (*Language Technologies LT*) ondoan, **hizkuntza-baliabideak** (*Language Resources LR*) jarri behar ditugu. Izan ere, azkenak ezinbestekoak dira aurrenekoak garatzeko. Hizkuntza-baliabideak gero eta garrantzitsuagoak izango dira, edozein hizkuntzatarako, denak izango baitira beharrezkoak zerbitzu eleanitzak inplementatzeko. Hortaz, corpusak (elebakarrak eta eleanitzak, parekatuak edo lerrotatuak), ahotsaren teknologietarako baliabideak... oso ondo kotizatuko dira, bereziki produktuak merkaturatu nahi dituzten enpresentzat.

Europar elkarte bat dago, ELRA (*European Language Resources Association*), hizkuntza guztietako baliabideak banatzen dituen: **www.elra.info**.

Hizkuntza-baliabideak direla eta, Europar Batasunak kezka du formatuen eta estandarren inguruan eta datozen urteotarako horiek harmonizatzeko Europar eredia finkatzea pentsatzen ari da, agian direktiba baten bidez. Lehentasunezko arloak ezartzea litzateke helburua, betiere kontuan izanda interesik handiena luketenak, eta, antza, horretan, ez litzukete aintzat hartuko bakar-bakarrik Europar Batasuneko hizkuntza ofizialak. Horrela, metodoak kontsolidatzeaz gain, *roadmap* bat egingo litzateke eta ebaluatzeko protokoloak ere finkatuko lirateke. Europar Batasunak orain gutxi liberatu du bere itzulpen-corpusa (itzulpen-memoriak), 6 milioi dokumentu, nahi duenak erabil dezan.

Hizkuntzen teknologiei dagokienez, teknologiak hor daude baina oraindik guztiak elkarrekin jartzeko azpiegitura bat falta da.

5. Euskararen erronkak.

Euskarak garapen teknolinguistikoa behar du. Tren hori hartzen ez duten hizkuntzek ez dute etorkizunik izango. Argi dago merkatuak horrelako produktuak eskatuko dituela eta dauden hizkuntzan kontsumituko dituela. Hemen, estrategia litzateke lehentasunak ezartzea, betiere, kontuan izanda gure merkatuan etorkizun hurbilean sartuko diren produktuak. Beraz, helburua litzateke produktu horiek euskaraz ere eskuragarri izatea. Arlo batzuk jorratu dira eta jorratzen ari dira baina lanean jarraitu beharko da. Aipatu bezala, giltzarria Internet (bai eta eskuko telefonoak ere) izango da eta, hori horrela, argi dago edukiak sortu beharko direla euskaraz. Honek izan behar du lehentasuna, web semantikoen teknologiak eduki bolumen handia eskatzen duelako. Dena den, aintzat izan behar da web eleanitzean oinarritzen dela teknologia hori.

6. Hizkuntzen teknologien gaineko bi helbide interesgarri.

<http://www.lt-world.org/> *Language Technologie world* (Alemania)

<http://www.lang-tech.org/> *LangTech is the european forum dedicated to communities and organisations involved in the development, deployment and exploitation of Language and Speech technologies.*

7. Bibliografia

- Aizpurua, Xabier; Diaz de Lezana, Araceli: "El euskera y las nuevas tecnologías de la información y de la comunicación: Programa de Industrias de la Lengua", in Segundo Congreso Europeo de Planificación Lingüística, Andorra la Vella, 2001.
- Diaz de Lezana, Araceli: "Euskara eta garapen teknolinguistikoa", www.erabili.com/zer_berri/muinetik/1057933133, 2003.
- Diaz de Lezana, Araceli: "El desarrollo tecnolingüístico del euskera", in Congreso Linguapax, Barcelona 2004.
- Hizkuntza politika aurrera begira. Hizkuntza Politikarako Sailburuordetzaren 2005-2009 egitasmoa, Eusko Jaurlaritzaren Argitalpen Zerbitzu Nagusia, Gasteiz, 2005, ISBN 84-457-2407-X.
- Langtech 2008, Hizkuntzen Teknologien gaineko Biltzarra. Erroma, 2008.
- Wikipedia, http://en.wikipedia.org/wiki/Main_Page