

# EUSKARAREN GARAPEN TEKNOLOGIAK

Araceli Díaz de Lezana  
Hizkuntza Politikarako Sailburuordetza  
Eusko Jaurlaritza

Bartzelona (Linguapax), 2003/05/21 (itzulpena)

## 1. Sarrera.

Informazioaren eta komunikazioaren aroan bizi gara. Aro honetan beharizan berriak sortu dira eta sortzen ari dira, neurri handi batean Internetek eraginda. Hori dela eta, unibertsitateak, ikerketa-zentroak eta enpresak hizkuntzen trataera berezia duten teknologia berriak garatzen saiatzen dira, "infoingeniaritza linguistikoa" edo "hizkuntzen teknologia" hain zuzen ere. Negoziobide eta merkatu berriak agertu dira, eta, horiek bultzatuta, teknologia horiek aurrerapauso handiak egiten ari dira; "hizkuntzaren industriak" legez ere ezagutzen dira, eta batik bat honako hauek dira:

- Ahotsaren ezagutza (**S**peech-**T**o-**T**ext y **A**utomatic **S**peech **R**ecognition)
- Ahotsaren sintesia (**T**ext-**T**o-**S**peech)
- Karaktereen ezagupen optikoa (**O**ptical **C**haracter **R**ecognition)
- Itzulpen automatikoa
- Ahozko datuen konprentzioa
- "Data / audio mining" (dokumentazio-multzo handietan bilaketak egitea)

Herritarrak informazioaren eta komunikazioaren gizartean zeharo integratzeko politikak jorratzen ari dira gobernuak eta administrazioak. Izar ere, konturatu egin baitira, batetik, Informazioaren eta Komunikazioen Teknologien garapenaren ondorioz gizarte-, ekonomia- eta politika-aldaketak sortzen ari direla, eta, bestetik, teknologiek erabiltzeak ahalmen ia mugagabea eskaintzen duela informazioa eskuratu, bildu, prozesatu eta partekatzeko. Hori horrela, Europako Batzordeak "eEuropa-Informazioaren Gizartea pertsona guztientzat" ekimena jarri zuen martxan 1999an, hainbat ekintza-planekin jarraipena izan duena.

## **2. Euskara eta informazioaren eta komunikazioaren teknologia berriak (IKT).**

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

Indarrean dauden hainbat planek bideratzen dute Eusko Jaurlaritzak Informazioaren eta Komunikazioaren Teknologia berrien eta euskararen arloan mamitu nahi duen politika:

## **√ 1999, Euskara Biziberritzeko Plan Nagusia.**

Plan honek –Euskararen Aholku Batzordeak egina, Eusko Jaurlaritzak onartua eta Eusko Legebiltzarrak berretsia– hurrengo 10 urteetan euskarak izango dituen erronka nagusiak ezartzen ditu, hiru mailatan sailkatuta:

1. Helburu nagusia: “euskaraz bizi nahi duten herritarrei horretarako aukerak eskaintzeko behar diren hizkuntza-politikako neurriak erabakitzea eta bultzatzea”.
2. Hiru helburu estrategiko: euskara ondorengoetaratzea; euskararen gizarte-erabilera areagotzea, euskarazko produktu eta zerbitzuen eskaintza zabalduta; eta euskararen kalitatea hobetzea.
3. Ia ehun neurri edo helburu zehatz, hamazazpi ekintza-sektore estrategikotan jorratzeko: planaren VI.3.2.c atalak “Informazioaren eta komunikazioaren teknologiak berriak” aipatzen ditu.

Bertan neurri hauek finkatzen dira:

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 oinarrizko tresnen azpiegitura sortu.

**√ 2001-2004, Zientzia, Tecnologia eta Berrikuntzarako Plana:**  
**Hizkuntza industriren programa.**

Programaren helburu nagusiak hauek dira:

1. Hizkuntzen Industriarekin zerikusia duten Enpresa eta Ikerketa Zentroen gaitasun zientifikoa eta eraberritze teknologikoa indartzea.
2. Enpresa- eta gizarte-sarean Hizkuntzen Industriaren sarrera bultzatzea.
3. EAEko aniztasun linguistikoaren eraginez sortutako eskakizunetatik eratorritako ezagutza teknologikoari etekina ateratzea.

Helburu horiek lortzeko, Planak lehentasunezko sei esparru hauek bultzatu nahi ditu:

- Hizkuntza-ingeniaritzaren oinarrizko tekniken garapena.
- Baliabide linguistikoaren garapena.
- Bitarteko aurreratuen erabilera bultzatzea eta plataforma gaitasun emaileen esperimendazioa.
- Informazioa bildu, berreskuratu eta bilatzea.
- Ahotsaren ezagutza automatikoa, ahotsaren konpresio digitala eta testua ahots bihurtzea.
- Testuaren tratamendurako funtzionalitate lagungarriak, bai testua ofimatikaren testuinguruan tratatzeko, baita testutik testurako aldaketan ere (itzulpena).

## ✓ 2002-2005, Euskadi Informazio Gizartean Plana

Plan honen helburua informazioaren teknologiak gizarte-maila guztietan sustatzea da, hain zuzen ere hezkuntza-alorrean, etxeetan, enpresetan, kulturean eta Administrazioan.

Planak 10 jarduera-ildo ditu:

- erabiltzaileei dagokienez: Internet guztiontzat, enpresa digitala eta Administrazioa on line
- eduki eta zerbitzuei dagokienez: Euskadi Sarean, e-Prestakuntza, e-Osasuna, edukien sektorea, eta
- testuinguruari dagokionez: teknologia eta araudia.

Plan hau oinarri hartuta, hainbat ekimen sortu dira, esaterako, "Konekta zaitez" eta "Enpresa Digitala". "Konekta zaitez" kanpainari esker, Euskadiko etxeen %34k ordenagailua Interneti konektatuta dute. Enpresa Digitala ekimena, berriz, enpresei zuzenduta dago, ekonomia berrian egokitzeko aholkularitza eskaintzeko.

Beste ekimen bat "KZgunea Proiektua" da; 2001ean jarri zen martxan, eta horren helburua da informazioaren eta komunikazioaren teknologia berriak erabiltzeko eta horietan prestatzeko berriazko hainbat gune publiko sortzea; bestalde, etxeko andreak, jubilatuak, langabetuak eta etorkinak teknologia berrietara erakarri nahi dira. Udalen, EUDELen eta Eusko Jaurlaritzaren artean sinatutako hainbat akordiori esker, egun 225 KZgune dira EAEn.

Ekimen horiezaz aparte, Euskadi Informazio Gizartean Planaren "Euskadi Sarean" atalean ere 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**

Arestian aipatutako ekimenetatik sortutako informazio-sistemen proiektuak eta beste batzuk hartzen ditu plan honek. Honako jardun-ildo hauek hartzen ditu aintzat:

- Barruko kudeaketa hobetzea
- Administrazio elektronikoa
- Azpiegiturak eta antolakuntza

### **3. Euskararen garapen teknolinguistikoa.**

Euskarak asko aurreratu du azken urteotan teknologia berrietan integratzeko:

1. Euskarazko prentsa on-line (euskara hutsezko egunkari bat eta izaera ezberdineko hainbat aldizkari).
2. Euskara ikasteko eta euskarazko edukiak dituzten CDak.
3. Diziplina anitzeko euskarazko web orriak, euskarazko bilatzaileak ([www.kaixo.com](http://www.kaixo.com), [www.aurki.com](http://www.aurki.com), [www.jalgi.com](http://www.jalgi.com)) eta gaikako aldizkari digitalak eta eztabaida-guneak ([www.sustatu.com](http://www.sustatu.com) eta [www.erabili.com](http://www.erabili.com)).
4. Sarean jarritako hiztegiak, corpusak eta hainbat erreminta, oso erabilgarriak hizkuntzaren teknohiztunentzat:

- XUXEN euskararen zuzentzaile ortografikoa, Euskal Herriko Unibertsitatearen Informatika Fakultatearen IXA taldeak egina. Doan jaitsi daiteke [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft) helbidetik: Office 97, 2000 eta XP bertsioak, bai eta OpenOffice.org-etik ere Windows zein Linux-erako ere; eta Macintosh-en Word 98 eta 2001 bertsioak.
- EUSKALTERM, Terminologia Banku Publikoa (UZEIk 1987an sortutako izen bereko bankuaren gainean sortua): [www.euskadi.net/euskalterm](http://www.euskadi.net/euskalterm)
- Hiztegi elektronikoak:
  - Elhuyar Hiztegia, gaztelania-euskara/euskara-gaztelania: [www.euskadi.net/hizt\\_el](http://www.euskadi.net/hizt_el)
  - 3000 Hiztegia, gaztelania-euskara/euskara-gaztelania: [www.euskadi.net/hizt\\_3000](http://www.euskadi.net/hizt_3000)
  - Sinonimoen Hiztegia, Antonimoduna: [www.euskadi.net/hizt\\_sinon](http://www.euskadi.net/hizt_sinon)
  - Morris Student Plus Hiztegia, ingelesa-euskara/euskara-ingelesa: [www.euskadi.net/morris](http://www.euskadi.net/morris)
  - Harluxet hiztegi entziklopedikoa: [www.euskadi.net/harluxet](http://www.euskadi.net/harluxet)
  - Elhuyar EU/ES hiztegiaren plugin-a; Word 2000 eta 2002an eta Internet Explorer-en integratzen da: [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft)
  - Euskaltzaindiaren Hiztegi Batua: euskararen ohiko lexikoaren 20000 hitz ingururen forma estandarra biltzen du, eta egun beste horrenbeste hitz sartzeko lanean ari dira: [www.euskaltzaindia.net](http://www.euskaltzaindia.net)

- XX. mendeko euskararen corpus estatistikoa; Euskaltzaindiak aginduta UZEIk egin du: [www.euskaracorpora.net](http://www.euskaracorpora.net)
- Euskararen Datu-Base Lexikala (EDBL): Euskal Herriko Unibertsitatearen Informatika Fakultatearen IXA taldeak burutua: <http://ixa2.si.ehu.es/edbl/>
- EAEko Toponimia-datutegia: [www.euskadi.net/euskara\\_eaetoponimia](http://www.euskadi.net/euskara_eaetoponimia)
- Euskal Autonomia Erkidegoko toki-entitate, euskal abizen (orain arte 10.000 inguru bildu dira) eta ponte-izenen izendegi ofizialak: [www.euskaltzaindia.net](http://www.euskaltzaindia.net)
- Euskal Herriko Agintaritzaren Aldizkaria (EHAA), euskaraz eta gaztelaniaz: [www.euskadi.net/bopv](http://www.euskadi.net/bopv)

5. Software produkturik erabilienean itzulpena eta lokalizazioa:

- Windows 95 eta Windows 98; Windows XP interfaze-aren paketea: [www.microsoft.com/spain/windowsxp/euskera](http://www.microsoft.com/spain/windowsxp/euskera)
- Word 6 eta Office 2000 (Word, Excel eta Outlook).
- Internet Explorer 4 eta 5; Outlook Express 5.
- Linux Mandrake sistema eragilearen 8.2 eta 10.0 bertsioak: [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft)
- OpenOffice.org 1.0.2 ofimatika-paketea, bai eta 1.1 bertsioa Windows eta Linux-erako: [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft)
- Panda Titanium Antibirusa: [www.pandasoftware.es/titanium/euskara](http://www.pandasoftware.es/titanium/euskara)
- Panda Platinum Antibirusa: [www.pandasoftware.es/platinum](http://www.pandasoftware.es/platinum)

- Autoikaskuntzarako Wit ikastaroak (Word 2000 eta Internet 2000).
- Omnipage-ren 11. eta 12. bertsioetarako OCR, bai eta Windows Word-erako eta Mac Word 2001erako ere: [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft)
- Optenet, Interneten edukietarako iragazkia: [www.optenet.com](http://www.optenet.com)
- Prisma II: Pyme-ak kudeatzeko softwarea
- SAPen bost kudeaketa-modulu: [www.euskadi.net/euskara\\_soft](http://www.euskadi.net/euskara_soft)

Horiezaz gain, eskola softwarea (EIMA) euskalduntzeko programak ere aipatu behar dira, Eusko Jaurlaritzaren Hezkuntza, Unibertsitate eta Ikerketa Sailekoak.

#### 6. Agertoki birtualak:

- HABENET: Helduen Alfabetatze eta Berreuskalduntzerako Erakundeak burututako proiektua, urrutiko irakaskuntza eta irakasleen prestakuntzarako. Hona helburuak: HABEren sarearen (HABE, euskaltegiak, irakasle eta ikasleak) barne- eta kanpo-harremanak hobetzea, material didaktikoa eskuragarri jartzea eta homologatzea, autoikaskuntza- eta didaktika sistemak on line garatzea; irakasleen eten gabeko prestakuntza tutoreen bidez gidatzea, administrazio-kudeaketa hobetu eta erraztea.
- HIZPONET, hizkuntza-plangintzaren kudeaketa optimizatzeko. Honen helburua da Euskararen gizarte-azpiegitura publiko eta pribatuetan barne- eta kanpo-komunikazioa hobetzea.

- HIZNET hizkuntza-plangintzari buruzko graduatuondoko ikastaroa da eta Eusko Ikaskuntzaren Asmoz Fundazioak, Udako Euskal Unibertsitateak (UEUK) eta Soziolinguistika Institutuak (SEIk) antolatu dute Hizkuntza Politikarako Sailburuordetzarekin batera. Ikastaroa Interneten bidez ematen da, eta hizkuntza-normalkuntzarako teknikariei, profesionalei, ikerlariei eta hizkuntza-normalizazioaren arloan ardura dutenei zuzentzen zaie.
- LINGUANET, prestakuntza-moduluak on line, eremu sozioekonomikoan diharduten hizkuntza-normalkuntzako teknikariei zuzenduak. Asmoz Fundazioaren eta Hizkuntza Politikarako Sailburuordetzaren ekimena da.
- Zinegotziek euskara ikasteko moduluak on line jarriko dituzte laster Hizkuntza Politikarako Sailburuordetzak eta Euskadiko Udalen Elkarteak (EUDELeK).

#### 7. Informazioa biltzea, berreskuratzea eta bilatzea:

- Kapsula softwarea: indexatu eta kontsultatzeko euskaraz merkaturatu den lehen softwarea; lematizazio-teknikak erabiltzen ditu (dokumentuak kudeatzeko sistema legez garatua). Proiektua Ametzagaina eta Adur Software enpresek garatu dute, eta Eusko Jaurlaritzaren Intek programaren bidez finantzatu da: [www.kapsula.com](http://www.kapsula.com)

#### 8. Ahotsaren ezagutza automatikoa eta sintesia:

- Telefónicak Euskal Herriko Unibertsitatearekin batera hizkuntza-gaitasunak jorrazteko hainbat proiektu garatu ditu, euskaraz ahotsa interpretatu eta soinu bidezko mezuen sintesia egiteko.

- Eusko Jaurlaritzak Scansoft Belgium enpresa belgikarrarekin lankidetzeta-hitzarmena sinatu zuen 2003an, batetik, ahotsaren ezagutza eta sintesirako motoreak sortzeko, eta, bestetik, euskarazko elkarrizketa kudeatzeko motoreak. Teknologia alorreko agente euskaldunak (enpresa eta unibertsitateak) proiektua bideratzeko behar diren hizkuntza-baliabideak egiten ari dira: datu-base akustiko bi, lexiko fonetiko bat eta testu-corpus handi bat. Horren helburua da ahots-teknologiak erabiltzen dituzten aplikazioak garatzen dituzten enpresek euskara integratzea produktuetan.

9. Hizkuntzaren ingeniartzaren oinarritzko teknikak garatzea, hala nola hizkuntza prozesatzeko teknikak. Euskal Herriko Unibertsitateko Informatika Fakultateko IXA taldeak ([www.ixa.si.ehu.es](http://www.ixa.si.ehu.es)) honako lan hauek burutu ditu:

- MORFEUS analizatzaile morfologikoa.
- EUSLEM, euskararako lematizatzaile/etiketatzaile automatikoa.
- Gain: euskarara egokitutako Internet/Intranet bilatzailea.
- Multimeteo: eguraldi-iragarpenak sortzeko sistema euskaraz.

Aipatzekoa da, halaber, Deustuko Unibertsitateko DELi ikerketa-taldea, informazio eta komunikazioko teknologietako lerroa irekita baitu: [www.deli.deustu.es](http://www.deli.deustu.es)

10. Testua testu bihurtzea (itzulpena) eta testuaren tratamendua testuinguru ofimatikoan; funtzionaltasunak:
  - Itzulpen-memoriak: lege- eta administrazio-itzulpenen memoriaren corpusa sortzen ari da Herri Ardularitzaren Euskal Erakundearen Itzultzaile Zerbitzu Ofiziala, Trados teknologiaren oinarriaren gainean.
  - Itzultzen laguntzeko tresnen zenbait prototipo, Euskal Herriko Unibertsitatearen IXA taldeak garatuak.
  - ATS teknologiaren esperimentazio-proba, euskarari aplikatuta; Eusko Jaurlaritzak eta Automatic Trans enpresa katalanak lankidetzan burutu dute.
  
11. Hizkuntza-ingeniaritza ikertzeko laguntza-programak:
  - ETORTEK programa (2002-2005), Eusko Jaurlaritzako Industria Sailarena: HIZKING 21, hizkuntza-ingeniaritza XXI mendearen atean, proiektua bideratzen ari dira euskal enpresa eta unibertsitateen eragile teknologiko elkartuak.
  - INTEK programa, Eusko Jaurlaritzako Industria Sailarena: programa honen barruan ACOTE proiektua garatu zen, eta horren oinarriaren gainean gaur egun AMETRA proiektua garatzen ari da. Proiektu biotan itzultzen laguntzeko tresnak ikertzen dira.
  
12. Hizkuntza-infoingeniaritzako enpresak sortzea: ELEKA enpresa; Adur Software, Ametzagaina eta CodeSyntax enpresak ere alor horretan ari dira.
  
13. Hizkuntza-teknologia irakaskuntzan sartzea: HIZTEK graduatuondoko ikastaroa, Udako Euskal Unibertsitateak antolatua.

14. Euskara Informazioaren eta Komunikazioaren Teknologia Berrietan integratzeko dirulaguntzak: helburu horretarako, urtero agindu bat argitaratzen du Eusko Jaurlaritzaren Hizkuntza Politikarako Sailburuordetzako Euskara Sustatzeko Zuzendaritzak. Era berean, helburu bera duten dirulaguntzak ematen dituzte Bizkaiko eta Gipuzkoako foru aldundiek. Horiezaz gain, ekimen lokalago batzuk ere badira.

#### **4. Ondorioak eta etorkizuneko erronkak.**

Argi dago, Informazioaren eta Komunikazioaren Teknologiei dagokienez, euskara ez dagoena hizkuntza nagusien parean. Dena den, azken urteotan euskarak aurrerapauso handiak egin ditu teknologia horietan integratzeko. Etorkizunari begira beharrezkoa izango da:

- Euskarak Interneten duen presentzia areagotzea, hurrei eta gazteei edukiak eta zerbitzuak eskainita batez ere.
- Interneten eskuragarri dauden tresnak eguneratu eta elikatzea, hala nola, hiztegiak, datu-baseak, corpusak, Terminologia Banku Publikoa, euskarazko softwareak; bestalde, tresna berriak ere sortu beharko dira, esaterako, gramatika- eta estilo-zuzentzailea (laster eskuragarri egotea espero dugu).
- Ahotsaren teknologietan oinarritutako aplikazioak garatzea, merkatuak hori eskatzen baitu: dei-zentroak (banku-eragiketak, mediku-hitzorduak, erreserbak, erosketak, trafiko- eta eguraldi-informazioa, iragarpenak...), dokumentu digitalizatuen kontsultak, irakasteko materialak; jokoak eta abar.

- Testuak bizkor ulertzen dituzten teknologiak garatzea: itzulpen automatikoa testutik testura eta ahotsetik ahotsera (speech-to-speech translation); testuak laburtzeko tresna baat; eta dokumentazio-multzo handietan bilaketak egitea eta ezagutza kudeatzea ahalbideratzen duen erreminta bat: informazioa indexatu, sailkatu, kodetu eta erauzteko edo husteko.

Ondorioz, premiak ezagunak dira, baina euskararen kasuan lehentasunak finkatu behar ditugu proiektuei ekin aurretik; izan ere, argi baitago ezin ditugula denak batera garatu.

Bestalde, kontuan izan behar dugu gure merkatua txikia dela, baina apurka-apurka eskaera handiarazi behar dugula, eta horretarako egon dauden produktuak ezagutarazi eta horien kontsumoa bultzatu behar da. Horrez gain, oso inportantea da lehendik gaztelaniaz, frantsesez edo ingelesez eskuragarri dauden aplikazioak euskaraz ere izatea.

Hizkuntza-teknologiak ez dira perfektuak. Hala ere, argi dago oso erabilgarriak eta lagungarriak direla gizarte honetan.

Euskara hizkuntza-teknologietan integratzeko, hizkuntza-baliabideak sortu behar dira, nahita ez. Garapen teknolinguistiko hori behar-beharrezkoa da bai hizkuntza arlo horretan normaliza dadin, bai euskararen presentzia bermatzeko gizartearen espazio elektroniko, telematiko, digital eta erretikularrean, bai eta globalizazio-garai honetan gure hizkuntzak bizirik iraun dezan ere.