

LAN-ARLOAK ETA -AZPIARLOAK

Euskarazko IKTen inbentarioaren aplikazioaren barnean, euskararen eta IKTen sektoreko eragileek gauzatzen dituzten proiektuen eta baliabideen informazioa katalogatzeko/biltzeko helburuarekin, **lan-arlo eta -azpiarlotan** egituratu da informazio hori:

➤ HIZKUNTZA-BALIABIDEAK ETA -TRESNAK

Hizkuntza Teknologietan oinarritutako aplikazioak garatzeko, funtsezkoak dira hizkuntza-baliabideak. Prozesaketa automatikoaren bidez hizkuntzaren ezagutza adierazteko modu nagusietako bat dira. Hauek dira hizkuntza-baliabideak:

- **Lexikoak, hiztegiak:** hainbat testuingurutan, egitura gramatikalari (morfologia), egitura fonetikoari (fonologia) edo esanahiari (semantika) buruzko informazioa ematen duen hitz-multzoa da.

Adibidez: hitzen arteko erlazioak (sinonimoak, antonimoak, izenak, multzoak, etab.) deskribatzen dituzten Hitz Sareak (Wordnets). Oso erabilgarriak dira informazioa, itzultzen laguntzeko tresna integratuak eta dokumentuak sortzeko sistema ofimatiko adimenduak bilatzeko.

- **Gramatika konputazionalak:** hizkuntza baten egitura deskribatzen dute, hainbat mailatan: hitza (morfologia), esaldia, perpausa, etab.

Adibidez: gramatikek esanahiaren (semantika eta diskurtsoa) edo sintaxiaren ikuspegitik lant dezakete egitura.

- **Corpusa:** idatzizko edo ahozko hizkuntzaren lagin handia da. Hizkuntza aztertzeko eta hizkuntzaren ezaugarriak zehazteko balio du, hizkuntzalaritzaren teoria enpirikoki egiaztatzeko eta makinak entrenatu eta haien jokaera egokitzeko.

Hizkuntzaren erabilera egokia definitzeko erreferentziazko corpusak, arlo jakinei buruzko corpusak (zientzia, teknologia...) eta corpus elebidun etiketatu eta lerrokatuak daude, oso erabilgarriak itzulpen automatikoko sistemetarako.

- **Beste zenbait baliabide:** ahots-baliabideak (datu-base akustikoak, lexikalak, fonetikoak...), thesaurusak –hierarkikoki elkarrekin erlazionatuta dauden terminoen zerrenda (askotan konplexuak dira), dokumentuak indexatzeko eta berreskuratze erabiltzen dena–, datu-base lexikalak...

Hauek dira, besteak beste, tresnak: zuzentzaile ortografikoa eta gramatikala, lematizatzaileak (oinarritutako tresna hizkuntza flexiboetarako), dokumentu eleanitzak editatzeko laguntza, analizatzaileak (morfologikoak, sintaktikoak edo semantikoak), etiketatzaileak (etiketak jartzen dizkiete hitzei, arau jakin batzuei jarraiki), testuak digitalizatzeko OCR sistemak (Optical Character Recognition), banku terminologikoak (jakintza-arloetako termino espezializatuei buruzko informazioa ematen dute), ontologiak (kontzeptuen deskribapen formala eta horien arteko erlazioak)...

➤ e-IKASKUNTZA

Internet bidezko autoikaskuntza da (hezkuntza birtuala). Ikaskuntza-erritmoa eta ikasteko tresnen erabilera ikasleari egokitzeko aukera ematen du, ordutegi- edo geografia-mugak baztertuta.

Hezkuntza-estrategia horrek hainbat tresna ditu: batetik, edukiak (testuak, animazioak, grafikoak, bideoak) aurkezteko zenbait erabilera, eta, bestetik, ikasleen arteko edo ikasle eta tutoreen arteko komunikazio-tresnak (posta elektronikoa, txata, foroak).

➤ ITZULPENAK EGITEN LAGUNTZEKO TRESNAK

Itzulpen-memoriak; itzulpen automatikoko aplikazioak.

➤ SISTEMA ELKARRERAGILE AURRERATUAK

Gure ingurunean IKTak sartzean, erabiltzaileen eta makinaren arteko erlazioa naturalagoa bihurtuko da. Hizkuntza Teknologiaei esker, pertsonen eta ingurunearen arteko elkarreragina lortuko da, bai lanean bai etxean; alegia, Ingurune Adimena gauzatuko da.

Komunikatzeko modu berri horri esker, aurrerapen handia egin da makinekiko elkarrekintza-tresna gisa hizkuntza naturalaren erabileran. Hala, hizkuntza ulertzeko eta sortzeko –idatzizkoa edo ahozkoa– sistemen arloan aurrerapen handia egingo da.

Hauek dira sistema elkarreragileak:

- **Ahotsa automatikoki ezagutzeko sistema (ASR, Automatic Speech Recognition):** horren bidez, ahozko mezu bat testu bihurtzen da. Gaur egun, pertsona batek makina bati hitz egiten badiu, makinak informazio hori ulertzeko gaitasuna du. Ordenagailuak uhin analogiko moduan jasotzen ditu ahozko soinuak, eta analizatu egiten dira hitzak osatzen dituzten unitateak (fonemak) identifikatzeko.

Adibidez: ordenagailu bati testu bat diktatzeko aukera dago; ordenagailuak idatziz jasoko du ahoz esandako guztia, teklatura erabili gabe.

- **Ahots-sintesia (TTS, Text-To-Speech):** horren bidez, idatzizko mezu bat ahozko bihurtzen da. Ahotsa sintetizatzekeo sistemak badaude, estilo eta emozio ugariak.

Adibidez: ikusmen-arazoak dituzten pertsonen ordenagailuaren pantailan duten informazioaren berri izan dezakete, idatzizko testua ahozko bihurtzen duten sistemen bidez.

- **Galderei erantzuteko sistemak (Question answering):** galdera bat aztertzekeo eta erantzun bat sortzekeo gai diren arrazonomendu-metodoak. Hainbat iturritako informazioa berreskuratzen da horretarako.

- **Ahots bidezko identifikazio biometrikoa:** esatariak ezagutzeko sistema da. Hitzuna nor den ezagutzen du edo egiaztatzen du automatikoki, ahots-seinalearen bidez. Besteena ez bezalakoa da gizaki bakoitzaren ahotsa. Horrenbestez, pertsona bat ahotsaren bidez identifika daiteke. Pertsonen ahots-identifikazioaren bidez, pertsona horrek zerbitzu bat erabiltzeko baimena duen edo ez egiaztatzeko teknikak garatu dira, besteak beste.

Adibidez: bankura transferentzia bat egiteko deitu eta sistema batek kasu egitea. Sistemarekin komunikatzen gara, eta hark ulertzen gaitu eta gure agindua betetzen du, baldin eta pertsona baimenduaren ahotsa bada.

➤ **INFORMAZIOAREN ETA EZAGUTZAREN KUDEAKETA**

Hizkuntza naturala da komunikaziorako eta ezagutzaren transmisiorako bide nagusia. Azken urteotan, esponentzialki handitu da on line argitaratutako informazioa –testuak edota multimedia–, gero eta nabarmenagoa da ezagutza hori guztia erabiltzeko eta ustiatzeko aukera ematen duten teknologien beharra.

Gaur egun, gero eta gehiago gordetzen da informazioa formatu digitalean, ordenagailu edo bestelako gailu elektronikoen bidez erabiltzeko. Hori dela eta, datu-baseetan, fitxategietan, web orrietan, data warehouse-etan, datu-biltegietan eta abarretan bilaketa eraginkorrak egiteko tresnak sortzen ari dira, balio erantsiko informazioa lortzeko, esanahia hartzeko, antolatzeko, erabilgarri bihurtzeko eta, azkenik, jakintza bihurtzeko.

Hizkuntza Teknologien aurrerapenei esker, laster on line dagoen informazioaz baliatzeko aukera unibertsala izango da, gero eta jende gehiagorentzat, teknologia menderatzen eta menderatzen ez dutenentzat. Aplikazio-eremu asko ditu horrek. Hauek dira interesgarrienak: iturri eta hizkuntza ugariko jakintza erauztea (dokumentazioa berreskuratzea, *Information retrieval*), dokumentazioa sailkatzea, dokumentazioa helbideratzea, dokumentazioa iragaztea, hizkuntza naturalean egindako galderei arrazoizko erantzuna ematea edo laburpenak eta testuak automatikoki egitea. Sistema horiek ez dute testu-informaziorako soilik balio; ikus-entzunezko informaziorako ere balio dute.

➤ **EUSKARAZKO SOFTWAREA**

Beste hizkuntza batean dagoen jatorrizko softwarea euskaraz lokalizatzea eta euskarara itzultzea ere sartzen da, baita euskarazko softwarea sortzea ere –librea edo jabeduna–. Erabilera askoko tresnak dira (sistema eragileak eta ofimatika-paketeak): enpresa-kudeaketa, salerosketa, irakaskuntza, hizkuntza, aisialdia, Interneteko tresnak, etab.

➤ **EDUKIEN INDUSTRIA**

Euskarazko eduki digitalak sortzea.

➤ **e-GIZARTERATZEA**

Teknologia erabiltzeko zailtasunak dituzten kolektiboak –minusbaliotasun fisikoak edo adimenekoak dituztenak– Informazioaren eta Komunikazioaren Gizartean

sartzeko tresnak dira. Gizartean Hizkuntza Teknologia erabiltzeak asko laguntzen die kolektibo horiei, baina, oro har, ez dira kontuan hartzen garapen teknologikoetan.

Arlo horretan, elementu hauek sartzen dira:

- Informazioaren Gizartean minusbaliatuak erabat integratzeko tresnak.
- Minusbaliatuen autonomia pertsonala garatzeko tresnak.
- Kolektibo guztiek teknologia erabiltzeko moduko ingurune birtualak.
- Giza talde baztertuak eta immigranteak gizarteratzeko ikaskuntza-plataformak, hizkuntzak eta kulturak ikas ditzaten.
- Pertsona guztiek bizi-kalitatea hobetzeko beharrezkoak diren sistema automatikoak erraz erabiltzeko interfaze adimendun eta atseginak.