

HITZARGIAK – Elkar argitzen duten hizkuntzak

ElkarOla

INFORMAZIO OROKORRA

DATA: 2016-04-29

1	Izena	Elhuyar
2	Hizkuntza	Euskara
3	Lan-arloa (hezkuntza, komunikazioa, irakaskuntza, aisia, administrazioa...)	IKT, hizkuntza-baliabideak
4	Kontakturako helbidea edo telefonoa	j.aztiria@elhuyar.com 943363040
5	Sare sozialetako kontaktuak (Webgunea, Facebook, Twitter...)	https://www.elhuyar.eus/eu/site/ikerketa-non

PRAKTIKA EGOKIAREN DESKRIBAPENA

1	Praktikaren azalpena	ElkarOla 2015-2016 urte-tarteetarako ikerketa estrategikoko proiektu bat da, euskararentzako oinarrizko baliabide teknologikoak sortzea, eta baliabide horietatik tresnak eta aplikazioak garatzea helburu duena, euskara aro digitalean osasuntsu egon dadin.
2	Praktikaren azalpenaren laburpena	<p>2002tik gaur egunera Elhuyar buru izan duten zenbait proiektu estrategiko eraman dira aurrera.</p> <p>Proiektu horien guztien helburu nagusia da hizkuntza-, ahots- eta multimedia-teknoologiak garatzea, itzulpenaren, irakaskuntzaren eta edukien sektoreek osatzen duten hizkuntzen industriaren sektore ekonomikoaren euskarri teknologiko izateko, mundu digital, interkonektatu eta eleaniztun honetan lehiakor izateko, eta hizkuntza ingurune horretan osasuntsu egoteko.</p> <p>Helburu horiek lortzeko, hizkuntzen industriako enpresa askorekin lehenasunezko harremanak izateaz gain, beharrezko masa kritikoa eta nazioarteko ospea duen zientzialari- eta ikertzaile-komunitate bat dute</p>

proiektuotan parte hartzen duten kideek.

Aplikazioei begira jarrita daude eta egon dira proiektuok. Oinarrizko ikerketa bazter utzi gabe, aplikazio esperimentalak aurkezten saiatzen dira, azken buruan itzulpenen, irakaskuntzaren edo edukien alorretan diharduten enpresek garatu eta produktu bihur ditzaten.

Alor desberdinak landu dira, eta euskarak une honetan dituen baliabide asko proiektu horien bidez lortu dira: zuzentzaile ortografikoak, corpusak, terminologia-erazketan egin diren lanak, bilatzaileak, ahotsaren sintesia, ahotsaren errekonozimendua...

Hauetako dira, zehazki, orain arte landu diren ikerketa-lerro garrantzitsuenak:

- Oinarrizko baliabideak: testu- edo ahots-corpusa, lexikoiak, hiztegiak, ontologiak, gramatika konputazionalak, analizatzaile morfosintaktikoa, ahots-ezagutzea, ahots-sintesia, elkarrizketa-sistemak...
- Itzulpenak: itzulpengintzaren sektorean erabil daitezkeen sistemak, itzulpen automatikoa, itzulpen-memoriak, ahots-ahots itzulpen-sistemak eta bikoizketa automatikoa kasu.
- Edukiak: edukien sektorea hobetzen lagundu dezaketen sistemak, hala nola, informazio-bilaketa (elebarkarra, eleaniztuna, semantikoa, multimedia...), informazio-erazketa, idazketan laguntzeko sistemak (zuzentzaileak, adibidez), ezagutzaren kudeaketa, galderei erantzuteko sistemak...
- Irakaskuntza: irakaskuntzaren eremuan erabiltzeko sistemak dira; adibidez, tutore pertsonalak, e-learning sistemak, ahoskera zuzentzeko sistemak, ariketen eta adibideen erakitzeko automatikoa...

ElkarOlaren epealdian, goian aipatutako alor horiek lantzen jarraitzeaz gain, beste alor batzuetarako aplikazioak sortzera ere pasatu gara, hainbat teknologia horretarako nahikoa helduta

		baitaude. Enpresa- eta industria-mundurako aplikazioak (onlineko ospearean kudeaketa, pertsona-makina interfazeak...) zein osasungintza alorrekoak (farmakozainketa, ezgaitasunak...) garatzen ari gara.
3	Aurrekariak (Inplementazioa burutzeko motibazioa, identifikaturiko beharrak)	2002tik hainbat ikerketa-proiektu estrategiko jarri dira martxan, euskararen garapen digitalerako. Lehenengo izan zen Hizking21; gero, etorri zen AnHitz; ondoren, Berbatek; atzetik, Ber2tek, eta azkena da ElkarOla. Azken urteetan, proiektu horiek Elhuyarren lidergopean egin dira, baina hainbat ikerketa-zentro eta unibertsitateekin batera gauzatu eta garatu dira.
4	Helburuak	Ikerketa-proiektu horien helburua izan da euskararentzako oinarrizko baliabideak sortzea, eta baliabide horietatik tresnak eta aplikazioak garatzea, euskara aro digitalean osasuntsu egon dadin.
5	Eragin esparru nagusia eta albokoak	Hizkuntzen industria, Hiztegiak, Idazketa, Itzulpengintza, Irakaskuntza, Bilatzaileak
6	Erakunde laguntzaileak eta horien lan-arloak	Elhuyar Fundazioak, Tecnalía Research & Innovation fundazioak, Vicomtech-IK4 teknologia-zentroak eta Euskal Herriko Unibertsitateko (UPV/EHU) IXA eta Aholab ikerketa-taldeek eratzen dute proiektua aurrera eramateko partzuergoa
7	Inplementazio-epea	2002tik gaur egun arte.

INFORMAZIO GEHIGARRIA PRAKTIKA EGOKIEN FITXA

8	Hedapena (eragiten dien pertsona kopurua, barne-kanpo hedapena...)	Euskal Herriko ehun enpresarialari eta euskarazko hizkuntza-komunitateari begira matrxan jarritako proiektua da.
9	Inbertsioa (dirutan, denboran, pertsonetan, bestelako baliabideak...)	50 ikertzaile aritu dira lanean
10	Lortutako emaitzak	Hauek dira ikerketa-lerro bakoitzean orain arte lortu diren emaitza eta produktu esanguratsuenetako batzuk:

Baliabide orokorrak: Proiektuko gainerako zereginetarako behar diren oinarrizko baliabideak konpilatu eta garatu nahi dira ikerketa-lerro honetan. Horien artean, hauek adibidez:

- Internet bidez corpusak lortzeko tresnak, elebakarrak zein eleaniztunak, orokorrak edo domeinukoak, konparagarriak edo paraleloak.
- Internet bidez lortutako era guztietako corpusak.
- Pibotatze bidezko hiztegigintza automatikoko tresnak.
- Corpus elebakar zein eleaniztunetatik, paralelo zein konparagarrietatik, terminologia erauzteko tresnak, hobetuak eta hizkuntza berriak dituztenak.
- Hiztegi elebakar edo eleaniztunak, orokorrak edo terminologikoak.
- Hiztegietan ezagutza semantikoa txertatzeko tresnak.
- Ontologiak eraikitzekeo tresnak, orokorrak nahiz domeinukoak, eskuz edo automatikoki lan egiten dutenak.
- Ontologia orokorrak eta domeinuzkoak.
- Mendekotasun sintaktikoen aztertzailak.
- Esaldi eta sintagmen mugak zehazteko sistemak.
- Analizatzaile semantikoak.
- Hizketa jarraitua ezagutzeko motorra.
- Hizketa-ezagutzaileak euskararako eta ingeleserako.
- Testua ahots bihurtzeko sistemen ebaluazio objektiboko teknikak.
- Ahotsak eraldatzeko sistemak.
- Ahots-segmentuak detektatzeko sistemak.
- Elkarrizketen txanda-aldaketak detektatzeko sistemak.
- Esatariak sailkatzekeo teknikak.
- Elkarrizketa-sistemak
- Biometria
- Ahots sintetikoaren detekzioa

Itzulpenak: Egun dauden itzulpen automatikoko sistemak hobetzen eta hizkuntza-bikote berrietara hedatzea da ikerketa-lerro honen helburua. Eginkizun hauei heldu zaie:

- Corpus paralelo handiak konpilatzea, itzulpen automatikoko sistema estatistikoen entrenamendurako
- Itzulpen automatikorako postedizio-tresna
- Itzulpen automatikorako aurreedizio-tresna
- Arauetan oinarritutako itzulpen automatikoko sistemak hobetzea
- Estatistiketan oinarritutako itzulpen automatikoko sistemak hobetzea
- Adibideetan oinarritutako itzulpen automatikoko sistemak hobetzea
- Sistema hibrido edo konbinatuak lortzea eta hobetzea
- Interlingua oinarritutako itzulpen automatikoko sistemak hobetzea
- Itzulpen automatikoa ebaluatzekeo irizpideak finkatzea
- Gaztelania-euskara itzulpen automatikoa hobetzea
- Euskara-gaztelania itzulpen automatikoa garatzea

- Ingelesa-euskara itzulpen automatikoa garatzea

Informazioaren kudeaketarekin zerikusia duen ikerketa-lerroan, bai informazioaren tratamenduari dagokionez, bai informazio hori berreskuratzeari dagokionez, eginkizun hauek jorratu dira:

- Entitate-izen bereziak (pertsonek, enpresenak, erakundeak...), leku-izenak, denbora-adierazpenak eta zenbaki-adierazpenak ezagutzeko tresnak
- Kontzeptuen eta entitateen arteko harremanak erazteko tresnak
- Korreferentziak identifikatzeko tresnak
- Dokumentuetako mapa kontzeptualak ateratzea
- Dokumentuen laburpen automatikoa
- Informazio-bilaketako nazioarteko lehiaketetan euskara sartzea (CLEF, TREC)
- QA, IR eta IERen inguruan antolatzen diren nazioarteko lehiaketetan parte hartzea
- RDFko ezagutza-baseen erauzketa-sistema eta horiei buruzko galderak erantzuteko sistema
- Informazio eleaniztunaren bilaketa-sistema
- Galderei erantzuteko sistema
- Metadatu semantikoen editorea
- Esatarien diarizazio- eta transkripzio-sistema
- Irudien analisi-sistema
- Multimedia dokumentuetan bilaketa semantiko egiteko sistema
- Gertakizunen erauzketa automatikoa
- Azpitolazio automatikoa
- Albisteen kontsumorako interfaze aurreratuak
- Kultur-edukietarako atzipen-sistema aurreratuak
- Bertsoen ahots-sintesi abestua

Irakaskuntza-prozesua errazteko baliabideak eta tresnak garatzea ardatz duen ikerketa-lerroan eginkizun hauei heldu zaie:

- Ahoskera ebaluatzeko sistemak
- Hizketa ezagutzearen bidez erantzun-zuzentasuna baloratzeko sistemak
- Testuen ebaluazio automatikorako metodoak, ebaluazioari laguntzeko.
- Plagioak detektatzeko metodoak
- Corpusetik abiatuta ikaskuntzarako baliabideak (ariketak, adibideak...) sortzen laguntzeko plataforma
- Ikasleari hizkuntzen ikaskuntzan laguntzeko plataforma (testu-idazketarako laguntzak): hiztegiak, corpusak, zuzentzaileak, itzulpen automatikoa, sorkuntza morfologikoa

Fabrikazio aurreratuaren alorrerako lan ildo hauek landu ditugu:

- Onlineko ospearen kudeaketa
- Bezeroen asetasuna sare sozialetan denbora errealean jarraitzeko behatokia
- Salmenta- eta komunikazio-prozesuen kudeaketa aurreratua
- Teleasistentzia bidezko mantenimendurako gailu janzgarriak
- Testuen itzulpen automatikoa denbora errealean gailu janzgarrietan

Azkenik, osasungintzaren alorrean arlo hauetarako aplikaziotan lan egin da:

- Farmakoen eragailu kaltegarrien analisia eta detekzioa txosten medikuetatik
- Ahotsen transformazioa ahotsen berreskuratzeari aplikatuta
- Ahots sintetiko pertsonalizatuak
- Asalduradun ahotsen edo ahots patologikoen ulergarritasunaren hobekuntza
- Dislexiadun pertsonentzako dokumentu irakurgailua
- Autismodun pertsonentzako irakaskuntza-aplikazioak

**11 Dokumentazioa
erreferentziako
materiala**

12 Aurkezpen bideoa <https://www.youtube.com/watch?v=1ZWIn0a10vg>

**13 Intereseko informazio
gehigarrik** http://berbatek.ber2tek.eus/eu/ikerketa_erroak