

47. PLATA proiektua

Proiektuaren izena	PLATA: Domeinuari egokitutako itzulpen automatikorako plataforma			
Proiektuaren laburpena (esaldi 1)	Plataforma bat garatzea, domeinu jakin batzuetara egokitutako itzulpen-sistema automatikoak sortzeko eta erabiltzeko			
Proiektuaren hasiera-data	2012	Proiektuaren amaiera-data	2017	
Erakunde nagusia edo koordinatzailea	Vicomtech-IK4			
Parte hartzen duten beste erakunde batzuk	Erakundea	Proiektuari egindako ekarpen nagusia		
	MondragónLingua	Plataformaren ustiaketa		
	MCC Telecom	Plataformaren integrazioa		
	Bartzelonako Unibertsitate Autonomia	Domeinura egokitutako sistemen kalitatea eta erabilgarritasuna ebaluatzea		
	EiTB	Erabiltzailea		
	Orona	Erabiltzailea		
	Fagor Arrasate	Erabiltzailea		
	Alecop	Erabiltzailea		
	Galbaian	Erabiltzailea		
	Geminys	Erabiltzailea		
	Danobat Group	Erabiltzailea		
Proiektuaren aurrekontua (milaka euro)	Urtea	Aurrekontua guztira	EAeren parte-hartzea	
	2012	318.757	318.757	
	2013	322.482	322.482	
	2014	316.319	316.319	
	2015	426.218	314.144,5	
	2016	517.947	337.388,4	
	2017	614.545	375.339,3	
EAeren parte-hartzearen finantzaketa-iturriak (mila euro)	Urtea	1. finantzaketa: 2020. urteari begira	2. finantzaketa: Proiektuaren bazkideak	Bestelako laguntza publikoak
	2012		234.027,65	84.729,35
	2013		230.696,16	91.785,84
	2014		239.008,47	77.310,53
	2015		326.929	99.289
	2016		400.604,25	117.342,75
	2017		475.435,45	139.109,55
Jardun-eremua	Lehentasunezko arlo estrategikoak <small>Markatu X batekin</small>			
	Fabrikazio aurreratua	Energia		Biosanataria
	Aukera-esparruak <small>Markatu X batekin</small>			
Elikadura	Hiri-habitata	Ekosistemak	Kulturaren eta sormenaren arloko	

				industria
				X

Proiektuaren deskribapen laburtua: helburu nagusiak eta garatu beharreko emaitzak, zer erronkari erantzuten dion, ekonomian eta gizartean izan dezakeen inpaktua, eta abar.

Itzulpen automatikoak hainbat abantaila erabakigarri eskaintzen ditu enpresentzat hil ala biziko diren sektoreetan, bizkor nazioartekotzen eta produktibitatea handitzen laguntzen duen aldetik. Zehazki, itzulpen automatikoa Europako merkatu digital bakarrerako osagai nagusietako bat da; izan ere, europarren % 43 inguru ez daude prest online zerbitzuak edo produktuak erosteko, non eta ez diren euren ama-hizkuntzan eskaintzen¹.

Itzulpen automatikoko soluzioak eraginkorrakoak dira domeinu mugatuetan; kasu horietan ematen dute kalitate eta erabilgarritasun handiena. Domeinu espezifikoetara egokitutako sistemak garatzeko, gaur egun nagusi den paradigma itzulpen automatiko estatistikoa da, domeinuaren datu propioetan oinarritutakoa. Bi alderdi horiek uztartzeak aukera bikaina ematen du itzulpen automatikoko teknologia integratzeko, enpresen jarduera garatzen laguntzeko hainbat industria eta sektoretan.

Proiektu honen bidez, itzulpen automatikoko sistemak sortzeko eta erabiltzeko konponbide eraginkorrak eta malguak garatuko dira, enpresa bakoitzaren jarduerekin definitutako domeinu espezifiko batera egokituta. Helburu hori lortzeko, beharrezkoa da, alde batetik, itzulpen automatikoaren alorreko erronka teknologikoei erantzutea, eta, bestetik, teknologiaren erabilgarritasuna, integrazioa eta erabilera menderatzea. Proiektuaren azken helburua enpresen beharretara erabat egokituta dagoen itzulpen-zerbitzua ematea da, babes teknologiko bikainarekin.

PROIEKTUAREN HELBURUAK

Helburu nagusia lortzeko, proiektuak tarteko helburu jakin batzuk bete behar ditu hainbat mailatan.

Lehenik eta behin, ezinbestekoa da domeinuan egokitzeko teknikan ikertzea itzulpen automatiko estatistikokoaren alorrean. Alderdi horren barruan sartzan da domeinu espezifikoaren datuen integrazio optimoa, automatikoki sortutako itzulpenak domeinuan baliaitutako hitz eta esamoldeen arabera moldatuta. Era berean, metodo optimoen garapena eskatzen du domeinu jakin batekoak ez diren datuak jasotzeko eta aukeratzeko, behar dugun domeinuan datu gutxiegi egonez gero, behar bezala erantzun ahal izateko. Alderdi horrek eskatzen du sarean askotariko datuak biltzeko metodoak taxutu beharra daudela, itzulpen-sistema optimoak garatu ahal izateko datu-multzoak lortzeko. Azkenik, ezinbestekoa da ikerketa-ahalegin gogorra egitea itzulpen automatikoko oinarritzko teknikei dagokienez, sistemok sortutako itzulpenen kalitatea eta erabilgarritasuna handitzen joateko.

Bigarrenik, komeni da itzulpen automatikoarekin lotutako alderdiak inguru egoki bakarrean gordeta edukitzeko plataforma bat sortzea, hainbat eratako domeinuetara egokitutako itzulpen-sistemak modu eraginkorrean garatzeko. Itzulpenak ezaugarri askotako hizkuntzen artean egiten direnez, itzulpen automatikoaren eremuak metodo eta osagai espezifikoak eskatzen dituzten prozesamenduak behar dira, itzulpenarako azken sistemak sortzeko eta erabiltzeko, hain justu. Besteak beste, plataformak hainbat zereginetarako osagaiak bildu beharko ditu: corpusak eta datuak hainbat formatutan tratatzeko osagaiak, hizkuntza- eta itzulpen-ereduak, hizkuntza anitzeko itzulpen-prozesuak eta dokumentuen kudeaketa. Azken

¹ http://ec.europa.eu/public_opinion/flash/fl_313_en.pdf

alde horri esker, zehazki, produktibitatea nabarmen irabaziko da, modu gardenean itzulitako dokumentuak sortuko baitira, jatorrizko egitura eta formatu bera errespetatuta baina testuen edukia modu automatikoan itzulita. Plataformak erantzun eraginkorra eman behar die merkatuaren eskariaren arabera sortutako eskariei, eta horren barruan sartzen da azken sistemen garapena posible egingo duen inguru konfiguragarri batean beharrezkoak diren osagaien integrazioa.

Hirugarren helburua erabilgarritasunarekin lotuta dago. Garatutako plataformak itzulpenen erabileraren benetako beharrezan erantzun behar die; beharrezan horiek hainbat agertoki bilduko dituzte: zuzeneko itzulpen automatikoaren erabilera zuzena, edota giza profesionalak zuzendutako eta baliozkotutako itzulpen automatikoak. Horretarako, ezinbestekoa da ahaleginaren zati bat bideratzea osagai jakin batzuk garatzera, eta, hala, itzulpen automatikoaren funtzionalitate jakin batzuk lortzera (hala nola postedizioa eta itzulpenak baliozkotzat jotzea). Hortaz, PLATA plataformak itzulpen-funtzionalitate konbinatuak eskuratzeko modua ipintzen du; hau da, profesionalak eskaintako hizkuntza-zerbitzuekin uztartzeko aukera, enpresa bakoitzaren beharrei erantzun malgua eskaintzeko.

Azken helburua domeinu espezifikotara egokitutako sistemen bidez lortutako hobekuntzen ebaluazio objektiboa da, Google Translate, Microsoft Bing eta tankera horretako sarbide askeko sistemen bidez lortutako itzulpenen kalitatea eta erabilgarritasuna alderatuta. Hizkuntzaren datu generikoetan oinarrituta egonik, sistema horiek ez dute domeinu espezifikoen datuekin garatutako sistemen kalitate bera ematen, eta ezinbestekoa da, hain justu, kalitate eta erabilgarritasunari dagozkion aldeak ebaluatzea, sistema egokituak erabilia produktibitatean zenbateraino irabazten den balioesteko. Alderdi honi esker, modu objektiboan zabalduko dira irabazi posibleak, garatutako plataforma baliatuta; hala, datu objektiboak emango dira PLATA plataformaren bidez proposatutako soluzioa hartuz gero lortuko liratekeen abantaila posibleei buruz.

PROIEKTUAREN ERRONKAK

Honako hauek dira proiektuaren erronka nagusiak:

- Itzulpen automatikoko sistemak sortzea, hainbat enpresak zehaztutako domeinuetara egokituta;
- Itzulpen automatiko estatistikoan domeinura egokitzea eskatzen duten metodoak optimizatzea;
- hizkuntza-teknologia barreiatu eta zatikatuak integratzea;
- itzuli beharreko edukien bolumena etengabe ari da handitzen;
- nazioartekotze-beharrezan handiak, hizkuntza ugari biltzen dituztenak (batez ere Europako merkatuan);
- datuen konfidentzialtasuna babestea;
- itzulpengintza automatikoa lehendik martxan dauden enpresa-prozesuetan txertatzea.

PROIEKTUAREN INPAKTUA

Aipatutako erronkei eta helburuei erantzunez, proiektuaren inpaktuak honako alderdi hauek biltzen ditu:

- itzulpengintza automatikoko teknologien kalitatea eta erabilgarritasuna handitzea;
- itzulpengintza automatikoko teknologiak gero eta gehiago baliatzea enpresa-prozesuetan;
- itzuli beharreko edukien bolumena etengabe handitzeari erantzun efizienteagoa ematea;
- itzultako datuak babestea eta prozesuan konfiantza handiagoa;
- enpresentzat merkatu berrietan sartzeko denborak murriztea, batez ere ETEentzat;
- enpresek lehiakortasun handiagoa izatea nazioartekotzeari eta hizkuntza anitzeko edukiak sortzeari dagokionez, hainbat modutan;
- Europako merkatu digital bakarrari erantzun efizientea ematea.