

**Hizkuntzen egun europarra, irailaren 26a**

## **QTLeap proiektuko teknologia hibridoa itzulpen automatiko hobe baterako**



Azken egunetan milaka errefuxiatu iritsi dira Europara. Integrazioa bermatzeko premiazko eginbeharretako bat hartu dituzten herrialdeko hizkuntza irakastea da. Eginkizun honen tamainak europar integrazioari ez zaiola behar bezala heldu gogorarazi beharko liguke: traba linguistikoak dira Europako muga egoskorrenak, merkataritza, komunikazioa eta benetako integrazioa oztopatzen dutenak.

Hizkuntzak ikastea ez da nolanahiko eginkizuna. Haurrek jolasean edo helduei entzunez ikasten dituzte bai ama-hizkuntza zein atzerriko hizkuntzak. Helduek, ordea, erregela gramatikalak eta hiztegia ahalegin handiz bakarrik ikas ditzaketela konturatzen dira.

Interneten atzigarri dauden itzulpen automatikoko sistema ezagunak bezalako programa informatikoek haurren antzera ikasten dituzte hizkuntzak. Ikasketa automatikoko metodoak erabiltzen dituzte testu elebidun erraldoietan (“Big Data”) hitz eta hitz segidak nola itzuli ikasteko probabilitate estatistikoak atereaz. Sistema hauek itzulpen bat ematen dute, testu originalaren arabera kalitate oso desberdinekoa izan daitekeena.

Hizkuntzak, baina, ez dira hitz segida hutsak. Egitura gramatikalak dituzte, bere subjektu, predikatu eta objektuekin, esaterako. Badaude informazio linguistikoak baliatzen duten sistemak. Haatik, orokorrean sortzen dituzten egiturak zurruneziak izaten dira testu orokorretarako eta, testuinguru zehatzetarako moldatu, eta hortxe erabiltzen dira. Adibidez, *La Vanguardia* egunkaria gauero itzultzen da katalanera honelako sistema batekin eta prentsara bidali aurretik zuzentzaile batek berrikusten du.

Interneten, ordenagailuentzako egituratutako Wikipediako bertsioen antzeko baliabideek Web Semantikoa deritzona eratzen dute eta jakintza semantikoa (hitzen esanahia) gordetzen dute. Baliabide hauek munduari buruzko jakintza egituratua kodetzen dute. Adibidez, “badakite” Paris Frantziako hiriburua dela, baina baita Texas-en Paris izeneko hiri bat dagoela ere.

QTLeap (qtleap.eu) proiektu europarraren barruan, hizkuntzaren prozesamenduan diharduen IXA ikertaldeak, Euskal Herriko Unibertsitatean, itzulpen automatikorako teknologia hibridoak garatzen ari da, metodo estatistikoak aberasten dituen Interneteko jakintza linguistikoarekin eta semantikoarekin, itzulpen hobeak lortzeko.

Lehen prototipoak itxaropentsuak dira baina “oraindik ikerketa mota honen hasieran baino ez gaude” dio Eneko Agirrek, IXA-n proiektuko koordinatzaileak. “Pertsonak itzultzera jartzen garenean, jatorrizko testuaren esanahia ulertu, eta esanahi hori itzultzen dugu. Askotan itzulpen libreak ateratzen zaizkigu, eta itzulpen zehatza egitea asko kostatzen zaigu. Egungo teknologiek, aldiz, itzulpen literalak egiteko joera handia dute, esanahiei baino azaleko hitzei erreparatuz. Hitzen eta esaldien esanahia

automatikoki atzitzea eta itzulpen automatikoaren prozesuan txertatzea erronka handia da gure alor honetan, eta IXA taldea gogotsu ari da horretan.”

Egun gure haurrek orain dela hogeit hamar urte Wikipedia gabe etxeko lanak egiteko nola moldatzen ginen galdetzen diguten bezala, baliteke gure bilobek galdetzea nola moldatu zen komunikatzeko, 2015. urtean, Europara bizitza hobea baten bila etorri zen jendea *omni-itzultzailea* izan aurretik.

Informazio gehigarria eskuratzeko eta harremanetarako: [qt Leap.eu](http://qt Leap.eu).