

Informatikatik hizkuntzara

DUELA gutxi zenbait egunkaritan argitaratu zen 'Metafisikatik zientzietara eta zientzietatik informatikara' izenburuko iritzi artikuluar buruz ohar batzuk egin nahi genituzke, kasu honetan ez informatikaraino heltzeko, baizik eta informatikaren mundutik daukagun ikuspuntua azaltzeko.

1) «... edozeinek daki hiztegi batek edo entziklopedia batek oso erraz har dezakeela euskarri informatikoa: hartarako mila pezetako batzuk baizik ez dira behar. Horregatik Grolier CD-ROMean edo disko optikoan erabiltzen dela esateak ez du ezertxo adierazi nahi. Edozein hiztegiarekin egin daiteke hori epe labur batean» esaten da artikuluan. Ehun milako batzuk edo milioiak beharko lirateke pertsona batek testu guztia teklateu beharko balu. Egia da, gaur egunean hori ez dela behar izaten inprimategi guztietan euskarri informatikoa erabiltzen baita. Baina hala ere, euskarri informatikoan testu huts bezala dagoen entziklopedia batek ez du inor liluratuko. Alde batetik, horrelako errepresentazio batean hitz-bilaketak errazagoak lirateke, gainera definizioetan dauden hitzen artean ere bila liteke. Baina bestetik, entziklopedietako hitz kopurua ikaragarri handia denez konputagailu baten testu ere luzeegi suertatuko litzateke bilaketa horiek burutzeko denbora. Dena dela, informatikan beste errepresentazio bereziago batzuk asmatu izan dira tamaina erraldoiko informazioaren tratamendu egokia egin ahal izateko. Entziklopedia bat horrelako modu batez eskaini ahal izateko azterketa, diseinu eta inplementaziorako diru eta denbora dezentente beharko lirateke.

2) «... Egiten ez bada arrazoi simple bategatik da: CD-ROMetik edozeinek kopia baitezake». Benetan problema interesgarria da kopiarena. Ez da gutxiesteko kontua, baina ez

dugu uste hori denik Britannica Entziklopedia CD-ROMean oraingo ez ikusteko arrazoi erabakiorra. Bestalde hortxe dugu Oxford English Dictionary ospetsua, bere 11 tomoetako edukia CD-ROM bakar batean amurkezen duela. Zenbait puntu aipatu behar da kopiatzeko erraztasunak direla eta: a) kopia paperezko entziklopedietan ere gertatzen da. b) Gaur egun badago paperezko entziklopedia euskarri informatikora kopiatzeko modurik, scanner eta OCR programen bitartez esate baterako. Beraz... kontuz ibili! c) Entziklopedia 'elektroniko' batetik testu zatiak kopiatzea erraza da; testu osoa kopiatzea, ordea, gaitza da; baina informazioa heltzeko eskaintzen dituen erraztasunak, programen bidez inplementatzen direnak, horiek ezingo dira eskuratu. Eta informazio handiak bilatu eta konparatzeko erraztasun horiek dira entziklopedia mota berri horien ekarpen nagusia.

3) «... bestetik askoz ere arosoago eta errazago baita liburua zuzenean kontsultatzea, kontsulta hauek telefono bidez 'on line' egitea baino. Merkeago eta lau aldiz lasterrago (Euskalterm-era deitzea aski da, froga egiteko)». Lehenengo eta behin, entziklopedia elektronikoa bat kontsultatzeak ez du telefono beharrik, CD-ROM teknologia, a.b., eskura edukiz gero. Bistan da epe laburrean edozein etxezuloan izango dugula horrelakorik. Beste kontu bat da urruneko datu-bankuak atzitu nahi izatea, adibidez Euskalterm kontsultatzea telefono bidez. Horrelakoetan guztiz ados gaude artikulua horretan esandakoaz, kontsulta sinpleez ari bagara eta hegoaldean egun daukagun azpiegitura telematikoa eskasa erabili behar bada. Hau da, Fisika Hiztegiaren hitz bakar bat, esate baterako 'fuerza', nola esaten den jakin nahi badugu, eta bulegoan hiztegi hori edukiz gero, gaur egunean askoz lehena-

go lortuko dugu paperezkoa joz; baina itzulpen bat egiterakoan hamar kontzeptu ezezagun aldez aurretik bildu baditugu eta guztiak bi edo hiru hiztegi desberdinetan bilatu nahi izanez gero, bada... orduan lehenago bukatuko dugu Euskaltermera deituz. Gainera datu-base terminologiko hori beti egunean dago.

4) «... Hizkuntza kontuetan informatikarekin liluratzen denak gutxi daki, seguraski, informatikaz edo hizkuntzaz. Edo bietaz». Begiak zabaldu baino ez da egin behar 'Lengoaia Naturalaren Prozesamenduan' edo 'Linguistika Konputazionalen' mugitzen diren merkatu-produktuak, proiektuak, kongresuak, laborategiak eta entpresak ikusteko. Merkatuan dauden produktuak ehundaka kontatzen dira, baina % 90 baino gehiago ingeleserako asmatu dira. Kontua ez da, beraz, informatikarekin liluratzea ala ez, baizik eta ikustea informatikaren bidetik hizkuntzaren erabilerarako lagungarri izan daitekeen hainbat eta hainbat tresna garatzea posible dela. Eta ez hori bakarrik, zeren linguistika konputazionalaren ikerrarloan eskakizun den formalizazioa onuragarria gertatzen baita hizkuntzaren beraren ikerketa teorikorako.

Bukatzeko esan behar dugu ikerkuntza arlo hau eronka erakargarria dela, orokorrean linguista eta informatikarientzat eta, konkretuki, euskara bezalako hizkuntza minorizatu baten laguntza teknikoak nondik nora aplikatu behar edo ahal diren bilatzen dutenentzat.

KEPA SARASOLA

EHU-KO INFORMATIKA FAKULTATEKO
LENGOIA NATURALAREN
PROZESAMENDUA IKERKUNTZA TALDEA

Artikulua beste hauek ere sinatzen dute: X. Arregi,
X. Artola, A. Diaz de Illarza, K. Gojenola, M.
Maritxalar eta E. Agirre.