

AHOTSAREN TRATAMENDUAN OINARRITUTAKO INFORMAZIO-ESKURATZE SISTEMENTZAKO DATUEN OPTIMIZAZIOA

Tesiaren egilea: Aitzol Ezeiza Ramos

Unibertsitatea: Euskal Herriko Unibertsitatea (UPV/EHU)

Saila: Sistemen Ingeniaritza eta Automatika

Tesi-zuzendaria: Miren Karmele López de Ipiña Peña

Tesiaren laburpena:

Tesi honen helburu nagusia da oinarriak ezartzea hizketan oinarritutako Informazioaren Eskuratzeko sistema automatikoak garatzeko, euskara eta inguruko hizkuntzetarako. Horretarako, albistegien edukiak indexatzeko erronka aukeratu da, bere zailtasuna eta konplexutasuna tratatzeko metodologia egokia proposatuz.

Zehatzago, Hizketako Dokumentuen Eskuratzeko landu da, AdiUp izeneko sistema elkarlanean garatuz, taldekide ikertzaileekin eta enprekin kolaboratuz. Horretarako, bai hizketaren ezagutze automatikorako bai testuen gainean lan egiteko tresna erabilgarriak garatu dira, besteak beste hizkuntzaren identifikaziorako tresnak.

Horretaz gain, oinarritzko Hizkuntza Baliabideak garatu behar izan dira, albistegien ataza egin ahal izateko. Hori lortzeko, oinarritzko baliabideak eskuratzeko gain, prozesatu dira, eta hori talde-lanean egin ahal izateko metodologia garatu da.

Azkenik, oinarritzko Hizketako Dokumentuen Eskuratzeko Sistema garatzeko datu-eskasiaren arazoa ere tratatu da. Honelako lanetan ezinbestekoa da datuetatik ahalik eta informazio gehien lortzea. Ondorioz, optimizazio-tekniketan sakondu da eta, tesiaren ekarpen nagusien artean, ezaugarrien erauzketan ahotsaren seinalearen ezaugarri ez-linealetara modu egokian moldatzeko teknikak erabiltzea proposatu da. Gainera, sistema ez-lineal eta kaotikoen teoriaraino oinarritzen den Dimentsio Fraktalaren ezaugarria Hizketaren Ezagutze Automatikorako sistema tradizioalean barneratzeko metodologia proposatu da, horretarako esperimentuak garatuz.