

# ADVANCED FUZZY MATCHING IN THE TRANSLATION OF EU TEXTS

Margita Šoštarić

## **Sažetak**

*Prevoditeljski je posao danas nezamisliv bez alata za strojno potpomognuto prevođenje. Budući da prijevodne memorije čine ključnu komponentu tih alata, s pravom se postavlja pitanje kako ih bolje iskoristiti da bi se olakšao i ubrzao prijevodni proces. Ovim se istraživanjem želi ispitati postoje li učinkovitije metode pronalaženja korisnih prijevodnih prijedloga od onih koje se trenutačno upotrebljavaju u komercijalnim alatima. Točnije, u radu se istražuje dovode li kompleksniji algoritmi i uključivanje jezičnog znanja u proces utvrđivanja sličnosti segmenata do zamjetnog napretka u kvaliteti rada sustava prijevodnih memorija. Eksperiment se izvodi sa sintaktičkim stablima dobivenim parsiranjem prijevodne memorije GU za pismeno prevođenje (DGT-TM), na kojima se potom isprobava niz konfiguracija za utvrđivanje sličnosti segmenata. Uz pojedinačne se metode ispituje i učinak njihove kombinacije pomoću regresijskog modela. Odabrani segmenti ocjenjuju se automatskom evaluacijom, na temelju korelacija i automatski izračunanih srednjih vrijednosti te ljudskom evaluacijom, na temelju korelacija izvedenih iz anketnih odgovora ispitanika. Konačni je cilj istraživanja utvrditi bi li trebalo razmisliti o implementaciji jednog od testiranih algoritama u rad postojećih komercijalnih alata kako bi se unaprijedio prijevodni proces.*