

Kazalo

Predgovor	5
I. JEZIČNE TEHNOLOGIJE	
1. Što su jezične tehnologije?	9
1.1. Jezične tehnologije	12
1.2. Jezične tehnologije prema Petom okvirnom programu Europske Unije	14
1.3. Prethodnici jezičnih tehnologija	18
1.3.1. Tourski manifest i jezične industrije	18
1.3.2. Jezično inženjerstvo	23
1.4. Nastavak u Šestom okvirnom programu	26
1.5. Sastavnice jezičnih tehnologija	27
1.5.1. Jezični resursi	28
1.5.2. Jezični alati	29
1.5.3. Komercijalni proizvodi	41
2. Jezične tehnologije za hrvatski jezik	45
2.1. Europske integracije i hrvatski jezik	45
2.2. Prisutnost hrvatskoga jezika u digitalnom obliku	46
2.3. Strateški položaj jezičnih tehnologija	50
2.4. Razvitak jezičnih tehnologija za hrvatski jezik	53
3. Jezične tehnologije i pretraživanje informacija	57
3.1. Morfološki »osjetljivi« upiti	61
3.2. Prepoznavanje naziva	63

3.3. Semantički »osjetljivi« upiti	66
3.4. Višejezični upiti	72

II. PRIMJENA JEZIČNIH TEHNOLOGIJA NA HRVATSKI JEZIK

4. Portal jezičnih tehnologija za hrvatski jezik	77
5. Hrvatski nacionalni korpus	87
5.1. 30–milijunski korpus (30m)	88
5.1.1. Raspon	88
5.1.2. Opseg	89
5.1.3. Struktura	89
5.2. Hrvatski elektronski tekstovni arhiv (HETA)	91
5.3. Zapis korpusa	92
5.4. Dostupnost HNK	96
5.5. Perspektive razvitka HNK	98
6. Hrvatski morfološki leksikon	101
6.1. Teorijske osnove	102
6.2. Struktura leksikona	104
6.3. Leksikon osnova	106
6.4. Flektivni leksikon	107
6.4.1. Lemarij	112
6.4.2. Popis nastavaka	113
6.4.3. Popis preoblika	114
6.4.4. Generiranje oblika	115
6.5. Uporaba i dostupnost leksikona	116
7. Morfosintaktičko označavanje hrvatskih korpusa	121
7.1. Korpus	125
7.2. Istopisnost	126
7.3. Obradba	128

8. Primjena jezičnih tehnologija pri prevođenju pravne stečevine EU na hrvatski jezik	135
8.1. Prevođenje pravne stečevine EU na hrvatski	135
8.2. Obilježavanje termina u izvornicima	136
8.3. Pronalaženje kandidata za termine	141
8.4. Perspektive	146
9. Zaključak	149
III. POJMOVNIK JEZIČNIH TEHNOLOGIJA	153
Popis pokrata	171
Literatura	175
Bilješka o autoru	187