

Bojana Dalbelo Bašić

Jure Mijić, dipl. ing.

Jan Šnajder

Inovativnost u strojnoj obradi teksta – iskustva s CADIAL projekta

U prvom dijelu predavanja ukratko će se predstaviti aktivnosti i projekti Laboratorija za tehnologije znanja Fakulteta elektrotehnike i računarstva. Područje interesa grupe jest dubinska analiza teksta i podataka (*data and text mining*), inteligentno pretraživanje informacija i automatsko crpljenje informacija iz teksta (*information extraction*) te obrada prirodnog jezika (*natural language processing*). U drugom dijelu predavanja predstaviti će se način rješavanja problema predobrade morfološki složenog hrvatskog jezika (morfološka normalizacija) koja je pogodna za postupke dubinske analize teksta i pretraživanja informacija. U trećem dijelu opisat će se i demonstrirati način rada tražilice CADIAL, tražilice pravnih propisa Republike Hrvatske na kojoj su implementirani postupci morfološke normalizacije hrvatskog jezika, pretraživanje automatski indeksiranih dokumenata na temelju pojmovnika EUROVOC i višejezično pretraživanje. Projekt CADIAL (www.cadial.org) nagrađen je u prosincu 2009. godine Premijerkinom web-nagradom i dva puta nagradama VID I e-novation (2007. i 2009.).

Kratke biografije predavača

Bojana Dalbelo Bašić redoviti je profesor na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu. Diplomirala je 1982. godine na Matematičkom odjelu PMF-a Sveučilišta u Zagrebu. Do 1986. godine radila je u Institutu građevinarstva Hrvatske, a od 1986. do 1990. godine radila je u tvrtki Iskra-Delta. Od 1990. do 1999. godine radila je na Šumarskom fakultetu Sveučilišta u Zagrebu. Magistrirala je 1993. i doktorirala 1997. godine na Fakultetu elektrotehnike i računarstva u Zagrebu. Od 2000. radi na Fakultetu elektrotehnike i računarstva. Nositelj je kolegija “Umjetna inteligencija“, “Neizrazito, evolucijsko i neuroračunarstvo” i “Strojno učenje“, a na doktorskom studiju sunositelj je kolegija „Otkrivanje znanja u skupovima podataka“ i „Statistički modeli za dubinsku analizu podataka“.

Voditelj je više međunarodnih i domaćih, znanstveno-istraživačkih i stručnih projekata iz dubinske analize podataka i teksta. Autorica je više od 70 znanstvenih i stručnih radova. Dobitnik je DAAD stipendije 2000. Godine. Bila je gost predavač na Sveučilištu u Rennesu, Francuska i Katoličkom Sveučilistu u Leuvenu, Belgija.

Područja znanstvenog i stručnog interesa Bojane Dalbelo Bašić su umjetna inteligencija, strojno učenje, dubinska analiza podataka i teksta, pretraživanje informacija i obrada prirodnog jezika.

Jure Mijić diplomirao je 2006. godine na Fakultetu elektrotehnike i računarstva s radom iz područja pretraživanja informacija u tekstnim podacima. Nakon diplomskog studija upisao je poslijediplomski studij na istom fakultetu, smjer računarstvo, područje umjetne inteligencije, točnije obrada prirodnog jezika i ekstrakcija informacija iz teksta. Od 2007. godine radi na FER-u kao suradnik na međunarodnom projektu CADIAL, u sklopu kojeg je razvio tražilicu CADIAL kao jedan od rezultata tog projekta. Bio je na znanstvenom usavršavanju od 3 mjeseca na Katoličkom Sveučilištu u Leuvenu, također u sklopu projekta CADIAL. Do sad je objavio 3 rada, od kojih je jedan u časopisu skupine B, a dva u zborniku radova s međunarodnog znanstvenog skupa.

Jan Šnajder diplomirao je 2002. godine na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu, te magistrirao je 2006. godine na istom fakultetu. Od 2006. godine zaposlen je na Zavodu za elektroniku, mikroelektroniku, računalne i inteligentne sustave Fakulteta elektrotehnike i računarstva kao znanstveni novak. Njegovi znanstveni interesi obuhvaćaju obradu prirodnog jezika, pretraživanje informacija, dubinsku analizu teksta te funkcijsko programiranje. Sudjelovao je u četiri domaća te dva međunarodna znanstvenoistraživačka projekta.

Objavio je šest znanstvenih radova u međunarodnim časopisima te deset radova na međunarodnim skupovima.