

1. Raziskovalna organizacija:

Univerza v Ljubljani, *Fakulteta za računalništvo in informatiko*

2. Ime in priimek mentorja:

Marko Bajec

3. Področje znanosti iz šifranta ARRS:

2.07.07

4. Kontaktni e-naslov mentorja:

marko.bajec@fri.uni-lj.si

5. Kratek opis programa usposabljanja:

Mladi raziskovalec se bo ukvarjal z raziskavami na področju obdelave velikih količin podatkov (ang. Big data). Dejstvo, da smo v samo zadnjih dveh letih generirali več kot 90% vseh podatkov, ki so bili kdajkoli generirani, je osupljivo. Danes podatki nastajajo vsepovsod, njihova količina pa je tako velika, da jih težko ali povsem nemogoče procesirati z uporabo obstoječih orodij, metod in pristopov. Izzivov je veliko, kako takšne velike količine podatkov zajemati, hraniti, obdelovati, iskati po njih, jih vizualizirati itn. Prav tako je velik interes, da rešitve najdemo, saj verjamemo, da se lahko z učinkovito obdelavo velikih količin podatkov dokopljemo do pomembnih informacij kot na primer ugotavljamo poslovne trende, napovedujemo elementarne nesreče, preprečujemo goljufije, gradimo pametna mesta, realiziramo učinkovito rabo energije itn.

Mladi raziskovalec se bo osredotočil predvsem na obdelavo podatkovnega toka v realnem ali skoraj realnem času (ang. data stream analysis). Delal bo na razvoju ekspertnega sistema, ki bi znal iz semantike podatkov (s pomočjo ontologij) razbrati pomen posameznih atributov in pametno izbrati parametre in najbolj primerne metode za analizo podatkovnega toka. Podrobno bo preučil znane algoritme za analizo podatkovnega toka ter predvsem njihovo parametrizacijo. Razvit sistem bo preskusil na različnih sintetičnih podatkih ter podatkih iz realnega sveta. Vključen bo v raziskovalne aktivnosti programske skupine "Vseprisotno računalništvo" (P2-0359) ter druge projekte laboratorija.