

## DIPLOMSKI ZADATAK br. 2499

Pristupnik: **Tomislav Kravaršćan (0036491218)**

Studij: Računarstvo

Profil: Računarska znanost

Mentor: izv. prof. dr. sc. Igor Mekterović

Zadatak: **Informacijski sustav za prikupljanje javno dostupnih podataka za zadani skup osoba**

### Opis zadatka:

Potrebno je napraviti informacijski sustav koji će prikupljati podatke iz zadanog skupa internetskih portala, prepoznati radi li se u dohvaćenim podacima o nekoj od osoba iz zadanog skupa, te na temelju toga izgraditi bazu podataka koji bi uključivala vremenski ustrojen katalog članaka koji se odnose na zadane osobe. Sustav treba biti automatiziran i raditi kontinuirano. Redom, potrebno je ostvariti sljedeće zadatke: napraviti podesiv sustav prikupljanja članaka sa zadanog skupa internetskih portala. U okviru tog zadatka potrebno je proučiti postojeće besplatno dostupne alate za prikupljanje podataka s Interneta te pomoću odabranog alata izgraditi podesiv i robustan automatizirani proces prikupljanja. Zatim je potrebno izgraditi klasifikator koji će, koristeći katalog zadanih osoba (npr. bivši studenti FER-a), prepoznati spominje li se neka od osoba u tekstu i, potencijalno, prepoznati firme povezane uz ime osobe. Donijeti ocjenu uspješnosti ovog postupka. Prikupljene podatke je potrebno pohranjivati u bazu podataka. Konačno, potrebno je izgraditi web-aplikaciju koja će omogućiti nadzor i podešavanje sustava te uvid u prikupljene podatke, s odgovarajućim vizualizacijama. Donijeti ocjenu ostvarenog prototipa te smjernice za budući razvoj.

Rok za predaju rada: 28. lipnja 2021.