

Tema 4: Klasterizacija bookmark-a na Diigo social bookmarking sajtu

Korišćenjem [Diigo API-a](#) obezbediti pristup podacima o bookmarks-ima zadatog korisnika ili grupe korisnika. Uraditi tematsku klasterizaciju prikupljenih bookmarks-a na osnovu tagova pridruženih bookmarks-ima i/ili sadržaja (teksta) bookmark-ovanih Web stranica. Omogućiti pretraživanje bookmarks-a (kroz odgovarajući korisnički interfejs) na osnovu formiranih klastera.

Klasterizacija Web stranica na osnovu njihovog sadržaja (tj. teksta) je lepo objašnjena u poglavlju 3 (*Discovering Groups*) knjige [Programming Collective Intelligence](#).

Klasterizacija se može uraditi primenom K-Means ili Expectation Maximization algoritma.

Projektni zadatak se radi u Java programskom jeziku, uz korišćenje odgovarajućih Java framework-a; npr. Weka za klasterizaciju; [GSON](#) za procesiranje JSON zapisa u kome Diigo API vraća rezultate upita.

Proširenje teme za Završni rad: Za proširenje teme u Završni rad, isprobati i druge metode klasterizacije, kao što je hijerarhijska klasterizacija (lepo objašnjena u 3. poglavlju knjige [Programming Collective Intelligence](#)) i dati poređenje rezultata različitih algoritama.