Tema 9: Preporuka knjiga zasnovana na atributima knjiga (content-based recommendation)

Potrebno je kreirati sistem preporuke koji na osnovu zadate knjige preporučuje druge knjige, koristeći (kao osnovu za preporuku) sličnost knjiga zasnovanu na njihovim atributima. Drugim rečima, na osnovu knjige koji je korisnik zadao, sistem bi trebalo da preporuči npr. 5 knjiga koje su najsličnije zadatoj knjizi; preporučene knjige bi trebalo da budu sortirane prema izračunatoj sličnosti za zadatom knjigom.

Za realizaciju ovog sistema, potrebno je prvo kreirati dataset o knjigama preuzimajući podatke o knjigama iz DBpedia[1] baze znanja. Posebno je značajan atribut koji povezuje knjigu sa autorom (dbo:author), atribut koji povezuju knjigu sa žanrom (dbp:genre), kao i atribut koji povezuje knjigu sa temama kojima se bavi tj. koje je opisuju (dct:subject). Kao izvor podataka o knjigama, može se koristiti i Google Knowledge Graph[2] baza znanja posredstvom novog Search API[3]; u tom slučaju atributi koji opisuju knjigu, definisani su Schema.org vokabularom (konkretno klasom schema:Book[4])

Ovako kreiran dataset bi se koristio za preporuku knjiga na osnovu pristupa opisanog u radu: Linked Open Data for content-based recommender systems[5]. Iako se ovaj rad bavi preporukom filmova, opisani pristup kreiranju preporuka je podjednako primenljiv i na knjige (kao i na druge kategorije stvari). U radu je posebno značajna sekcija 2, s obzirom da ona opisuje model za predstavljanje podataka o entitetima i objašnjava način računanja sličnosti entiteta koristeći vrednosti njihovih atributa.

[1] http://dbpedia.org/
[2] https://googleblog.blogspot.rs/2012/05/introducing-knowledge-graph-things-not.html
[3] https://developers.google.com/knowledge-graph/
[4] https://schema.org/Book
[5] http://sisinflab.poliba.it/semantic-expert-finding/papers/tech-report-1-2012.pdf