

Društveno pretraživanje

Saik, Andrea

Master's thesis / Diplomski rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, University of Zagreb, Faculty of Humanities and Social Sciences / Sveučilište u Zagrebu, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:131:200943>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-14**



Sveučilište u Zagrebu
Filozofski fakultet
University of Zagreb
Faculty of Humanities
and Social Sciences

Repository / Repozitorij:

[ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences](#)



SVEUČILIŠTE U ZAGREBU
FILOZOFSKI FAKULTET
ODSJEK ZA INFORMACIJSKE I KOMUNIKACIJSKE ZNANOSTI
SMJER BIBLIOTEKARSTVO
Ak. god. 2018./ 2019.

Andrea Saik

Društveno pretraživanje: primjena, mogućnosti i izazovi

Diplomski rad

Mentorica : prof. dr. sc. Sonja Špiranec

Zagreb, rujan 2019.

Izjava o akademskoj čestitosti

Izjavljujem i svojim potpisom potvrđujem da je ovaj rad rezultat mog vlastitog rada koji se temelji na istraživanjima te objavljenom i citiranoj literaturi. Izjavljujem da nijedan dio rada nije napisan na nedozvoljen način, odnosno da je prepisan iz necitiranog rada, te da nijedan dio rada ne krši bilo čija autorska prava. Također izjavljujem da nijedan dio rada nije korišten za bilo koji drugi rad u bilo kojoj drugoj visokoškolskoj, znanstvenoj ili obrazovnoj ustanovi.

(potpis)

Sadržaj

1. Uvod.....	1
2. Pretraživanje informacija.....	2
3. Ključni aspekti pretraživanja informacija	4
4. Modeli pretraživanja informacija	6
5. Algoritmi za rangiranje.....	9
6. Kako je internet promijenio pretraživanje informacija?.....	10
7. Pretraživanje informacija na webu.....	12
8. Vrste upita na webu	14
9. Društveno pretraživanje.....	15
10. Aspekti društvenog pretraživanja	18
11. Implementacija društvenog pretraživanja	19
11.1. Korisničko označivanje	19
11.2. Sustavi pitanja i odgovora	22
11.3. Zajedničko pretraživanje	26
11.4. Suradničko filtriranje i sustavi preporuke	29
11.5. Personalizirano pretraživanje	33
12. Istraživanja društvenog pretraživanja	36
13. Zaključak	40
14. Literatura	42

1. Uvod

Društveno pretraživanje širok je pojam koji označava svaku vrstu pretraživanja koja omogućava korisnicima da pronalaze informacije na webu pomoću korisnički generiranog sadržaja. Cilj je ovog rada predstaviti mogućnosti, primjenu, izazove, probleme i ograničenja društvenog pretraživanja u svakodnevnom korištenju, odnosno njegovu upotrebljivost. Temeljno pitanje kojim se rad vodi jest može li se društveno pretraživanje rabiti kao efektivna nadopuna tradicionalnom obliku pretraživanja.

Kako bi se društveno pretraživanje uopće moglo razmotriti i usporediti s tradicionalnim, prvo poglavlje rada objasnit će pojam pretraživanja informacija i dati kratak prikaz razvoja ove discipline. U sljedećem poglavlju objasnit će se pojmovi relevantnosti, odziva i preciznosti, koji su ključni u pretraživanju informacija i evaluaciji procesa pretraživanja. Nakon toga bit će prikazani modeli pretraživanja informacija, kako su se razvijali od najjednostavnijih i statičnih, prema složenijima i dinamičnima, te kako su se odnosili prema korisniku i njegovim potrebama. Sljedeće poglavlje bavit će se algoritmima za rangiranje, točnije PageRankom, inovacijom koju su donijeli u proces pretraživanja i načinom na koji olakšavaju korisnicima pronalazak relevantnih rezultata. Također će biti naznačeni i problemi algoritma koji utječu na rangiranje rezultata i čine ga podložnim manipulaciji. Zatim će se navesti glavni utjecaji koje je nastanak i razvoj World Wide Weba imao na pretraživanje informacija, kao i određeni problemi koji su vezani uz pretraživanje na webu. U sljedećem će dijelu rada biti prikazana podjela korisnika prema stručnosti te vrste korisničkih upita i potrebe koje stoje iza njih.

U središnjem će dijelu rada biti opisana načela na kojima društveno pretraživanje počiva i navesti će se nekoliko različitih definicija društvenog pretraživanja kako bi se pokazale široke mogućnosti ovog pojma. Potom će se razmotriti nekoliko ključnih aspekata društvenog pretraživanja, kao što su vrijeme i oblik suradnje među korisnicima. Nakon toga će kroz pet različitih primjera biti prikazana primjena društvenog pretraživanja u različitim sustavima, uključujući korisničko označivanje, sustave pitanja i odgovora, zajedničko pretraživanje, suradničko filtriranje i sustave preporuke, te personalizirano pretraživanje. Opisat će se rad pojedinog sustava, mogućnost uporabe, te njihove pozitivne i negativne strane. U posljednjem poglavlju izložit će se rezultati nekoliko relevantnih istraživanja kako bi se prikazalo korištenje društvenog pretraživanja u svakodnevnom sluzenju internetom, zadovoljstvo korisnika takvim pretraživanjem i njegova uspješnost.

2. Pretraživanje informacija

Pretraživanje informacija (engl. *information retrieval – IR*) glavno je područje istraživanja informacijske znanosti, a procesi koji se odvijaju tijekom pretraživanja nastoje se proučiti i razumjeti kako bi bilo moguće konstruirati sustave pretraživanja koji olakšavaju učinkovitu komunikaciju i prijenos informacija između sustava i korisnika.¹ Hrvatska enciklopedija definira pretraživanje informacija kao „... traženje i izdvajanje nekoga podatka iz velikoga skupa pohranjenih podataka, npr. iz teksta, proračunske tablice, baze podataka, zbirke, kataloga, kazala, sekundarnoga dokumenta. Kod računalnih sustava provodi se zadavanjem upita (engl. *query*), obično jedne ključne riječi ili više njih, sintagme ili sl. Jedna od najvažnijih primjena danas je na internetu, gdje se na osnovi ključnih riječi pretražuju dokumenti u obliku mrežnih stranica.“² Iako danas pretraživanje informacija vežemo uz internet i računalnu tehnologiju, ustvari je riječ o disciplina koja postoji dugi niz godina. Čovječanstvo je počelo organizirati informacije za kasniju upotrebu i pretraživanje prije otprilike 4000 godina. Sakupljali su, organizirali i indeksirali glinene pločice, svitke papirusa te knjige.³ Najraniji računalni sustavi pretraživanja stvoreni su krajem 1940-ih, a bili su nadahnuti pionirskim inovacijama prve polovice 20. stoljeća.⁴ Sam termin skovao je Calvin Mooers 1950-ih godina, a jednu od definicija pretraživanja informacija dao je Gerard Salton, pionir u području pretraživanju informacija i jedna od vodećih ličnosti od 1960-ih do 1990-ih. Definirao je pretraživanje informacija kao područje koje se bavi strukturom, analizom, organizacijom, pohranom i pretraživanjem informacija. Iako je od nastanka ove definicije tehnologija značajno napredovala, ona je i dalje točna i odgovarajuća jer je termin „informacija“ jako širok i općenit, a pretraživanje informacija obuhvaća različite tipove informacija i aplikacija povezanih s pretraživanjem.⁵

Jedne od prvih institucija koje su primjenjivale sustave za pretraživanje informacija bile su knjižnice⁶, a upravljanje velikim zbirkama informacija potječe upravo iz

¹ Ingwersen, P. *Information retrieval interaction*. London : Taylor Graham, 1992., str. 49.

² Pretraživanje informacija. // Hrvatska enciklopedija. Mrežno izd. Zagreb : Leksikografski zavod Miroslav Krleža, cop. 2019. Dostupno na: <http://www.enciklopedija.hr/natuknica.aspx?id=50252> (16.8.2019.).

³ Ceri, S. ... et. al. *Web information retrieval*. New York : Springer, 2013., str. 3.

⁴ Sanderson, M.; Croft, W. B. *The history of information retrieval research*. // *Proceedings of the IEEE 100 (Special Centennial Issue)*, (2012), str. 1444.

⁵ Croft, W. B.; Metzler, D.; Strohman, T. *Search engines: Information retrieval in practice*. Reading : Addison-Wesley, 2010., str. 1.

⁶ Baeza-Yates, R.; Ribeiro-Neto, B. *Modern information retrieval*. 2nd ed. Harlow : Pearson, 2011, str. 3.

bibliotekarstva. Mogućnosti sustava za pretraživanje rasle su s povećanjem brzine procesora i kapaciteta pohrane, a ovaj je napredak također omogućio pomak od manualnog indeksiranja i pretraživanja u knjižnicama prema automatiziranim metodama.⁷ Napretkom tehnologije ovo je područje preraslo svoje prvotne ciljeve indeksiranja teksta i traženja korisnih dokumenata u zbirci dokumenata. Danas istraživanja u području pretraživanja informacija uključuju pretraživanje weba, modeliranje, klasifikaciju teksta, arhitekturu sustava, korisnička sučelja, vizualizaciju podataka, filtriranje, itd.⁸ Ključan korak u području pretraživanja označio je nastanak World Wide Weba donijevši trajnu promjenu u načinu pohrane, pristupa i pretraživanja zbirke dokumenata, te omogućavajući pristup javnosti i indeksirajući dokumente za sveobuhvatno pretraživanje.⁹ Time je pretraživanje informacija dobilo novo usmjerenje i istraživački fokus. Općenito govoreći, područje se može proučavati s dva prilično različita, ali komplementarna gledišta – prvo je usmjereno na sustav, a drugo na korisnika. Gledište koje je usredotočeno na sustav bavi se izgradnjom učinkovitih indeksa, obrađivanjem korisničkih upita s visokom uspješnošću i razvojem algoritama za rangiranje kako bi se poboljšali rezultati pretraživanja. S druge strane, pristup usmjeren na korisnika proučava ponašanje korisnika, njegove potrebe te kako te potrebe utječu na organizaciju i funkcioniranje sustava za pretraživanje.¹⁰

Primaran fokus područja još je od 1950-ih na tekstu i tekstualnim dokumentima poput web stranica, članaka, znanstvenih radova, knjiga, elektroničkih poruka i sl. Svi su ovi dokumenti karakterizirani strukturom koja može obuhvaćati naslov, ime autora, datum, sažetak, itd.¹¹ Informacije koje se pretražuju najčešće imaju oblik teksta, što znači da pretraživanje informacija odgovora pretraživanju teksta ili dokumenata, bilo da je riječ o cjelovitim tekstovima, broječanim podacima, bibliografskim informacijama i sl.¹² Osim pretraživanja tekstualnih dokumenata, pretraživati se mogu i multimedijски dokumenti kao što su slike, video i audio zapisi. Ove je medijske dokumente teško opisati i pretraživati, a trenutna tehnologija za pretraživanje ne-tekstualnih dokumenata češće se bazira na tekstualnim opisima sadržaja, a ne na samom sadržaju.¹³ Informacije se obično pretražuje u zbirkama nestrukturiranih ili polustrukturiranih podataka (npr. web stranice, dokumenti, slike,

⁷ Sanderson, M.; Croft, W. B. Nav. dj., str. 1444.

⁸ Baeza-Yates, R.; Ribeiro-Neto, B. Nav. dj., str. 1.

⁹ Ceri, S. ... et. al. Nav. dj., str. 3.

¹⁰ Baeza-Yates, R.; Ribeiro-Neto, B. Nav. dj., str. 1.

¹¹ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 1-2.

¹² Ingwersen, P. Nav. dj., str. 49.

¹³ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 2-3.

video, itd.). Potreba za sustavom pretraživanja javlja se kada zbirka dokumenata preraste opseg s kojim se tradicionalne tehnike katalogiziranje ne mogu nositi.¹⁴

Pretraživanje koje se temelji na upitu korisnika (ponekad se naziva i ad hoc pretraživanje jer je raspon mogućih upita širok i nije unaprijed određen) nije jedini tekstualni zadatak koji se proučava u pretraživanju informacija. Ostali zadaci uključuju filtriranje, klasifikaciju i odgovaranje na pitanja. Filtriranje se odnosi na otkrivanje zanimljivih priča na temelju interesa određenog korisnika i pružanja upozorenja putem e-pošte ili nekog drugog mehanizma. Klasifikacija ili organizacija rabi definirani skup oznaka ili klasa i automatski dodjeljuje te oznake dokumentima. Odgovaranje na pitanje nalikuje pretraživanju, ali je usmjereno na određena pitanja, npr. „Koliko su kilometara udaljeni Zagreb i Split?“. Cilj je ponuditi korisniku konkretan odgovor, a ne popis dokumenata.¹⁵

Iako je pretraživanje interneta najčešća primjena pretraživanja informacija, ono se primjenjuje i u korporacijama, vladi i drugim područjima. Vertikalno pretraživanje (engl. *vertical search*) posebna je vrsta pretraživanja weba gdje je područje pretraživanja ograničeno na određenu temu. Pretraživanje u poduzeću (engl. *enterprise search*) odnosi se na pronalaženje tražene informacije među mnoštvom računalnih dokumenata koji se nalaze na korporativnom intranetu. U ovom slučaju dokumenti obuhvaćaju web stranice, ali i izvore poput elektroničke pošte, izvještaja, prezentacija, tablica i baza podataka. Pretraživanje radne površine (engl. *desktop search*) individualna je verzija pretraživanja u poduzeću u kojoj se izvori informacija nalaze na pojedinačnom računalu, uključujući poruke elektroničke pošte i nedavno pregledane web stranice.¹⁶

3. Ključni aspekti pretraživanja informacija

Jedan od ključnih koncepata, ali i problema u pretraživanju informacija jest relevantnost. Široko gledano, relevantnost označava u kojoj mjeri pronađeni dokument odgovara upitu kojeg je korisnik postavio. Na korisnikovu prosudbu da li je određeni dokument relevantan utječu različiti faktori i moraju biti uzeti u obzir pri dizajniranju algoritma. Sama usporedba teksta upita i teksta dokumenta te traganje za egzaktnom podudarnošću obično ne daje relevantne rezultate. Jedan od razloga je taj što prirodnim

¹⁴ Sanderson, M.; Croft, W. B. Nav. dj., str. 1444.

¹⁵ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 3.

¹⁶ Isto, str. 3.

jezikom isti koncept može biti izražen na mnogo različitih načina i različitim pojmovima, a to dovodi do problema dvosmislenosti i nejasnosti rezultata.¹⁷

Relevantnost je koncept određen s nekoliko karakteristika. Kao prvo, relevantnost je subjektivna. Dvoje korisnika koji imaju istu informacijsku potrebu, mogu različito procijeniti isti dokument. Nadalje, relevantnost je dinamična u prostoru i vremenu. Dokumenti koje je sustav pronašao mogu utjecati na korisnikovu procjenu relevantnosti budućih dokumenata. Osim toga, korisnik može različito procijeniti isti dokument (dobiven temeljem istog upita), ovisno o njegovom trenutačnom stanju. Na kraju, relevantnost je višeznačna jer je određena ne samo sadržajem pronađenog dokumenta, već i vjerodostojnošću, iscrpnosti, jasnoćom i legitimitetom izvora. Bitno je napomenuti da relevantnost nije poznata sustavu prije korisnikove procjene.¹⁸ Učinkovitost pretraživanja informacija daleko je od sto postotne učinkovitosti. Korisnik koji traži informaciju, sam sustav i informacijski stručnjak ne mogu znati, i nikada neće znati, koje informacije nisu pronađene. Informacije koje sustav nije pronašao mogu se eksplicitno odnositi na upit, ali nisu pronađene zbog određenog nedostatka (npr. odgovarajućeg tezaurusa), ili mogu biti izostavljene jer potencijalno predstavljaju informaciju. To znači da su informacije implicitno sadržane u dokumentu, odnosno da ih korisnik iz njega može iščitati.¹⁹

Važno je razlikovati tematsku relevantnost (engl. *topical relevance*) i korisničku relevantnost (engl. *user relevance*). Dokument je tematski relevantan u odnosu na upit ako govori o istoj temi, a pri procjeni korisničke relevantnosti u obzir se uzimaju i drugi aspekti te se ona ne mora poklapati s tematskom relevantnošću. Primjerice, članak o pobjedi Kolinde Grabar-Kitarović nad Ivom Josipovićem na predsjedničkim izborima tematski je relevantan u odnosu na upit „predsjednički izbori“. Međutim, ako je korisnik koji je postavio upit već pročitao taj članak ranije, ili ako je zainteresiran za nadolazeće izbore, a ne one iz 2014. godine, za njega ovaj dokument neće biti relevantan.²⁰

Sljedeći ključan pojam u pretraživanju informacija je evaluacija. Kako bi se odredilo da li, i u kojoj mjeri, dokument odgovara potrebama korisnika, bilo je potrebno razviti načine evaluacije. Dvije mjere koje se koriste danas, preciznost i odziv, razvio je još u 1960-ima Cyril Cleverdon. Preciznost odgovara na pitanje koliko je pronađenih dokumenata relevantno, dok odziv mjeri koliko je relevantnih dokumenata u sustavu pronađeno. Kada se koristi mjera

¹⁷ Isto, str. 4.

¹⁸ Ceri, S. ... et. al. Nav. dj., str. 4.

¹⁹ Ingwersen, P. Nav. dj., str. 54.

²⁰ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 5.

odziva, obično se pretpostavlja da su svi relevantni dokumenti za određen upit poznati. Ova je pretpostavka problematična u web okruženju i mnoštvu potencijalno relevantnih dokumenata, dok može biti korisna u eksperimentima u području pretraživanja informacija. Takvi se eksperimenti sastoje od manje zbirke dokumenata, uzorka tipičnih upita i popisa relevantnih dokumenata za svaki upit.²¹

Na preciznost i odziv sustava utječu polisemija (postojanja više značenja za jednu riječ) i sinonimija (postojanja različitih riječi istog značenja), odnosno tzv. problem vokabulara (engl. *vocabulary problem*). Ovaj fenomen dovodi do nepodudaranja između upita i pronađenih dokumenata, jer je malo vjerojatno da će se pomoću dvije do tri ključne riječi uspjeti dohvatiti pravi dokument od milijuna dostupnih. Zbog sinonimije relevantne informacije neće biti uključene, jer upit neće sadržavati sve riječi kojima se može izraziti određen pojam, već samo jednu, što smanjuje odziv. Polisemija uzrokuje dohvaćanje nerelevantnih dokumenata jer će sustav pronaći sve dokumente koje sadrže upisanu ključnu riječ, a te ključne riječi mogu imati drukčije značenje od onog koje zanima korisnika. Ovo smanjuje preciznost sustava.²²

4. Modeli pretraživanja informacija

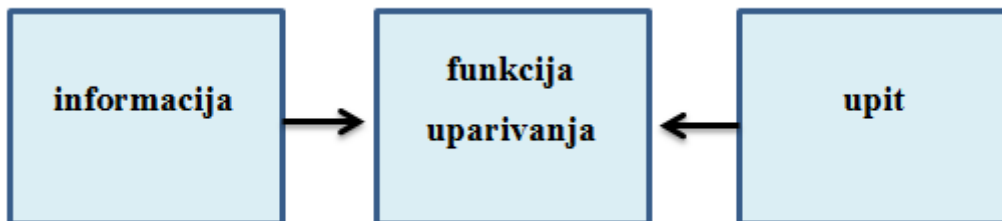
Modeli procesa pretraživanja razvijali su se tijekom godina i kretali od jednostavnih i statičnih do kompleksnih i dinamičnih. Klasični model jedan je od prvih modela pretraživanja informacija, a rabio se više od trideset godina u istraživanjima u području informacijskih znanosti, te je prikazan na slici 1. Bazira se na aktivnosti pretraživača koji uparuje informacijsku potrebu izraženu u upitu s dokumentima.²³ Lijevo se nalazi potencijalna informacija koja može stajati samostalno ili je predstavljena indeksnim pojmovima, grafičkim strukturama, kodovima kategorija (engl. *category codes*) ili formalnim podacima. Desno se nalazi informacijska potreba korisnika koja je predstavljena upitom i izražena prirodnim jezikom. U sredini se nalazi funkcija uparivanja koja uspoređuje informaciju ili njen prikaz s upitom i pronalazi dokumente ili dijelove dokumenta koji sadrže informaciju za kojom korisnik traga. Iz toga slijedi da je glavni cilj i problem pronaći dokumente koji će najbolje

²¹ Isto, str. 6.

²² Micarelli, A. ... et al. Personalized search on the world wide web. // The adaptive web : Methods and Strategies of Web Personalization / edited by Brusilovsky, P. et. al. Berlin : Springer, 2007., str. 198.

²³ Russell-Rose, T.; Tate, T. Designing the search experience: The information architecture of discovery. Amsterdam : Morgan Kaufmann, 2013., str. 24.

zadovoljiti informacijsku potrebu korisnika.²⁴ Iako je ovaj model imao znatan utjecaj, previdio je važnu stavku u procesu traženja, a to je korisnik. Kako bi pretraživanje bilo uspješno, treba postojati model koji u središte stavlja korisnika, a ne sistem.²⁵



Slika 1. Klasični model pretraživanja informacija²⁶

S druge strane, standardni model prepoznaje veću ulogu korisnika, a prema pretraživanju se odnosi kao prema tipu rješavanja problema. Ovaj model karakterizira niz četiri aktivnosti: identificiranje problema, artikuliranje informacijske potrebe, formuliranje upita, evaluacija rezultata. Korisnik prvo mora znati kakav problem ili zadatak ima, zatim izražava informacijsku potrebu u verbalnom obliku kao upit. Nakon toga pretraživač uparuje upit s dokumentima u zbirci, i dostavlja korisniku rezultate koje on potom evaluira. Ako je potrebno, može reformulirati ili precizirati upit. Ciklus se ponavlja sve dok korisnik ne zadovolji informacijsku potrebu.²⁷

Kao i standardni model, kognitivni model također gleda na pretraživanje kao na rješavanje problema potaknuto eksplicitnim ciljem korisnika. U ovome modelu, korisnik primjenjuje mentalni model kako bi stvorio plan djelovanja u svrhu ostvarivanja svog cilja. Mentalni model označava unutarnji prikaz problematične situacije i njenog konteksta. Razlika između kognitivnog modela i prijašnjih modela jest u tome što ovaj model prepoznaje bitnu ulogu znanja koje korisnik posjeduje o temi koju želi istražiti. Što je veća količina znanja, veća je i vjerojatnost da će korisnik uspješno oblikovati upit i procijeniti dobivene rezultate.²⁸

²⁴ Ingwersen, P. Nav. dj., str. 50.

²⁵ Russell-Rose, T.; Tate, T. Nav. dj., str. 24.

²⁶ Ingwersen, P. Nav. dj., str. 50.

²⁷ Isto, str. 24.

²⁸ Isto, str. 24.

Navedeni modeli dijele pretpostavku da informacijska potreba korisnika ostaje nepromijenjena tijekom procesa pretraživanja. Na proces pretraživanja gledaju kao na kontinuirano rafiniranje postavljenog upita sve dok korisnik ne pronađe idealan skup rezultata. Međutim, brojna su istraživanja otkrila kako informacijske potrebe korisnika evoluiraju u interakciji s informacijama, a korisnici stječu novo znanje pritom stvarajući i nove ciljeve. Prema ovom shvaćanju, pretraživanje je dinamičan, interaktivan proces u kojemu ono što pronađemo može promijeniti ono što tražimo.²⁹ Dinamički model koji odgovara ovim karakteristikama predložila je Batesova, uspoređujući pretraživanje informacija s branjem bobica (engl. *berrypicking*). Korisnici proces pretraživanja mogu započeti istražujući samo jednu značajku šire teme i nastaviti se kretati kroz različite izvore. Svaki novi podatak s kojim se korisnik susretne potencijalno može potaknuti nove ideje i smjerove, a s time se mijenja i upit. Ovo ne znači samo reformulaciju ključnih riječi kako bi se pronašao bolji rezultat, već je sam upit podložan stalnim promjenama. Informacijska potreba nije zadovoljena jednim konačnim skupom rezultata, već nizom izbora pojedinačnih referenci i dijelova informacija u svakoj fazi konstantno promjenjivog pretraživanja. Opisani proces analogan je branju bobica jer su bobice raširene na različitim grmovima i potrebno ih je brati jednu po jednu, nije moguće obrati cijeli grm odjednom. Upravo tako i korisnik pretraživanjem pronalazi jednu po jednu informaciju.³⁰

Jedan od modela koji je izgrađen na temeljima dinamičkog modela je model informacijskog putovanja (engl. *information journey model*). Glavne aktivnosti u ovom modelu su: prepoznavanje informacijske potrebe, stjecanje informacije, interpretiranje i validiranje informacije, upotreba informacije. Iako se ovi koraci čine slični standardnom modelu, zapravo su bliži dinamičnom modelu jer je naglasak na validaciji, interpretaciji i upotrebi informacije kao ključnim aktivnostima koje oblikuju evoluciju informacijske potrebe.³¹ U prethodnim modelima pretpostavlja se da je traženje informacija aktivni proces koji pokreće svjesno prepoznavanje eksplicitne potrebe. Međutim, postoje slučajevi kada korisnici dolaze do informacija bez da su ih aktivno tražili, bilo slučajnim susretom, neočekivanim uvidom ili sretnim otkrićem. Ti susreti s informacijama ono su što obično

²⁹ Isto, str. 25.

³⁰ Bates, M. J. The design of browsing and berrypicking techniques for the online search interface. // Online review 13, 5(1989), str. 410.

³¹ Russell-Rose, T.; Tate, T. Nav. dj., str. 26.

označavamo pojmom *serendipity* (događaj sretnog slučajnog otkrivanja), a ovaj model prepoznaje *serendipity* kao dio iskustva pretraživanja informacija.³²

5. Algoritmi za rangiranje

Rane su se tražilice oslanjale na tehnologiju pretraživanja informacija koja postoji od 1970-ih. Ukratko, rani pretraživači konstruirali su svoje vlastite indekse, indeksirajući stranice na webu i analizirajući sadržaj svake stranice, bilježeći riječi i njihove frekvencije na svakoj stranici. Kada je korisnik postavio upit, tražilica je dohvaćala i rangirala stranice koje sadrže pojmove upita. U počecima weba, naglasak je pri pretraživanju bio na veličini indeksa, a tražilice koje su indeksirale više sadržaja bile su u prednosti nad konkurentima. Pozornost je također pridavana rangiranju rezultata pretraživanja, a tražilice su se uglavnom oslanjale na učestalost pojmova upita na određenoj stranici kao glavnog načina za prosuđivanje relevantnosti. Iako je ovaj pristup funkcionirao razmjerno dobro u strukturiranom, zatvorenom svijetu sustava za pretraživanje informacija, gdje se od informacijskih stručnjaka moglo očekivati da će njihovi upiti biti detaljni i dobro oblikovani, nije imao isti uspjeh kada ga se primijenilo na širok i raznolik web sadržaj. Većina korisnika nije mogla biti zadovoljna takvim iskustvom pretraživanja, a relevantni rezultati ostajali su nedostupni. Poboljšanje rangiranja rezultata pretraživanja postalo je ključni izazov za rane pretraživače, a napredak se dogodio sredinom 1990-ih. Za njega su najzaslužniji osnivači Googlea Larry Page i Sergey Brin, te informatički stručnjak Jon Kleinberg.³³

Njihova se inovacija sastojala u tome da su iskoristili povezanost web stranica, te su te informacije koristili za procjenu relativne važnosti pojedinačne stranice. Na web su gledali kao na mrežu citata, a umjesto jednog rada koji citira drugi rad, na webu je jedna stranica citirala drugu stranicu kroz hipervezu koja ih povezuje. Ovo je pružilo osnove za razvoj novog načina vrednovanja važnosti web stranica i u kasnim 1990-ima algoritam PageRank implementiran je u Google.³⁴ PageRank funkcionira tako da stranica ima veću važnost i bolje je rangirana ako veći broj drugih stranica dijeli poveznicu na nju, pri čemu je bitan i autoritet stranice s koje poveznice dolaze, a ne samo njihov broj. Nakon Googleove inovacije, mnoge druge tražilice prihvatile su slične tehnike za rangiranje web sadržaja.³⁵

³² Isto, str. 27.

³³ Smyth, B.; Coyle, M.; Briggs, P. Communities, collaboration, and recommender systems in personalized web search. // Recommender Systems Handbook / edited by Rokach, L. et. al. Boston : Springer, 2011., str. 581-582.

³⁴ Isto, str. 583.

³⁵ Richman, D. Social search comes of age. // Information Outlook 11, 8(2007), str. 4.

Ipak, rangiranje rezultata nije savršeno i podložno je manipulaciji, a algoritmi imaju nekoliko osnovnih nedostataka. Kako bi većina korisnika bila zadovoljna pretraživanjem, PageRank osigurava da najviše rangirani rezultati budu i najpopularniji, odnosno ono što zanima većinu. Stranice koje nisu toliko popularne nisko su rangirane i mogu ostati neotkrivene i zanemarene, što znači da se nikada neće popeti više na rang ljestvici. PageRank pogoduje stranicama koje imaju veliku ekonomsku moć jer zbog svojeg marketinga i opće istaknutosti imaju puno poveznica. Na ovaj se način ojačava i podržava postojeća struktura moći jer popularnije stranice dobivaju bolji položaj što dovodi do većeg uvažavanja i izdvojenosti njihovih poruka i stavova. Osim toga, pretraživači ne indeksiraju sav web sadržaj u potpunosti, neke su stranice izostavljene ili su indeksirane samo djelomično.³⁶

Korisnici vjeruju kako su tražilice objektivne i vjerodostojne jer su osnovne operacije u njima automatizirane, uključujući i procese koje tražilice koriste za agregiranje svojih baza podataka i rangiranje tih podataka. Kako su rezultati pretraživanja stvoreni računalnim putem, a ne ljudskom odlukom, smatra se da tražilice ne mogu biti podložne predrasudama te da podaci u njima ne mogu prikazivati iskrivljenu stvarnost. Međutim, iako je potreba ljudske prosudbe u rangiranju velike količine pojmova uklonjena, i dalje postoji potreba za ljudskom uredničkom prosudbom u tom procesu. Odluka o tome koji će faktori biti uključeni u algoritme za rangiranje i koju će težinu imati, odražava ljudsku prosudbu o tome što čini neki sadržaj vrijednim. Sve navedeno utječe na stvaranje fenomena pod nazivom pristranost pretraživača (engl. *search engine bias*) koji opisuje sustavno favoriziranje određenog sadržaja od strane pretraživača, kako bi se zadovoljili korisnici.³⁷

6. Kako je internet promijenio pretraživanje informacija?

Područje pretraživanja informacija dugo je vremena bilo rezervirano samo za informacijske stručnjake i knjižničare, ali sve se promijenilo početkom 1990-ih nastankom World Wide Weba. Tim Berners-Lee 1989. pokrenuo je World Wide Web, a 1991. stvorio je prvu web stranicu. Internet je ubrzo postao najvećim repozitorijem ljudskog znanja u povijesti, zahvaljujući standardnom korisničkom sučelju koje je uvijek isto, neovisno o računalnoj okolini, i koje omogućava korisnicima da stvaraju vlastite dokumente. Posljedično,

³⁶ Goldman, E. Search engine bias and the demise of search engine utopianism. // Web Search : Multidisciplinary Perspectives / edited by Spink, A.; Zimmer, M. Berlin : Springer, 2008., str. 123-126.

³⁷ Isto, str. 122-123.

milijuni korisnika stvorili su milijarde dokumenta koji sačinjavaju znanje pohranjeno na internetu.³⁸ Najpoznatija primjena pretraživanja informacija i povezanih tehnika upravo je pretraživanje weba, a direktna je posljedica bitan utjecaj koji je web imao na razvoj pretraživanja informacija. Kao prvo, web je utjecao na karakteristike zbirke dokumenata. Zbirku dokumenata na webu sačinjavaju stranice koje se nalaze na preko milijun mrežnih mjesta, a povezane su poveznicama. Kako je zbirka dokumenata na webu distribuirana na različitim mjestima, prije indeksiranja potrebno je sakupiti sve dokumente i pohraniti njihove kopije u centralnom repozitoriju. Ova nova faza u pretraživanju informacija naziva se *web crawling*. Web pauci ili web roboti su softveri koji automatski preuzimaju web stranice, a te se stranice potom indeksiraju kako bi bile pretražive.³⁹

Sljedeći bitan utjecaj weba odnosi se na veličinu zbirke i opseg upita koje korisnici svakodnevno postavljaju pretraživačima. Kako je web je rastao više i brže nego bilo koja prije poznata zbirka dokumenata, a s njime i svakodnevni korisnički upiti, performansi i skalabilnost (mogućnost sustava da se nosi s rastućim brojem korisnika i dokumenata) pretraživača postale su ključne karakteristike sustava za pretraživanje. Nadalje, velika zbirka dokumenata otežava predviđanje relevantnosti. Korisnički upit obično pronalazi veoma mnogo dokumenata koji odgovaraju izrazima u upitu, ali velik broj tih dokumenata zapravo nije relevantan za korisnike.⁴⁰

Sljedeća dva utjecaja weba na pretraživanje proizlaze iz činjenice da internet nije više samo repozitorij dokumenata i podataka, već i mjesto poslovanja. Problem pretraživanja proširen je s traženja tekstualne informacije na druge korisničke potrebe poput cijena, brojeva telefona, poveznica za preuzimanje softvera i sl. Kako bi ove informacijske potrebe mogle biti adekvatno zadovoljene potrebna je identifikacija strukturiranih podataka povezanih sa željenim objektom interesa, poput cijene, lokacije, opisa nekih ključnih karakteristika. Posljednji se utjecaj odnosi na ekonomske aktivnosti koje se odvijaju na webu, poput oglašavanja i e-trgovine. Ove su aktivnosti često povezane sa spamom, a njegova prisutnost na webu još je dodatno otežala već ionako težak zadatak pronalaska relevantnih dokumenata.⁴¹ Jedna od uobičajenih metoda spama uključuje umetanje dodatnih ključnih riječi, ili čak cijelih rječnika, u nevidljivi tekst na web stranici kako bi stranica bila povezana s mnogo više korisničkih upita, iako nije relevantna za te upite. Ipak, kako se algoritmi za

³⁸ Baeza-Yates, R.; Ribeiro-Neto, B. Nav. dj., str. 3.

³⁹ Isto, str. 10.

⁴⁰ Isto, str. 11.

⁴¹ Isto, str. 11.

rangiranje baziraju na broju poveznica koje stranica dijeli s drugim stranicama, ovakve metode spama imaju ograničenu uspješnost. Međutim, *spammeri* su našli način kako da zaobiđu ovu prepreku tako da kreiraju mnoštvo web stranica koje sadrže poveznicu na stranicu čiju poziciju u rangiranju rezultata žele povisiti.⁴²

7. Pretraživanje informacija na webu

Internet je prvi izvor kojem će se većina ljudi obratiti kada ima neki problem ili pitanje, te kada žele saznati ili naučiti više o nekoj temi. Međutim, iskustvo pretraživanja nije uvijek ugodno i uspješno, a korisnici često ne pronalaze odgovor na svoj upit. Jedan je od čestih uzroka ovog problema nejasno izražavanje informacijske potrebe. To se može dogoditi jer korisnik ne zna kako svoju potrebu oblikovati u upit, ima praznine u znanju, ali i zato što korisnici često oblikuju kratke upite od dvije ili tri riječi. Na to su potaknuti prijašnjim iskustvima u kojima dugi upiti nisu donijeli rezultate. Istraživanja pokazuju kako su najčešći korisnički upiti kratki i sastoje se od dvije ili tri ključne riječi. Kratki upiti mogu biti dvosmisleni i neprecizni, a korisnici moraju uložiti trud kako bi izdvojili dvije ili tri riječi koje najbolje opisuju njihovu cjelokupnu potrebu. Čak i kada bi korisnici oblikovali upit pomoću velike količine pojmova, većina web pretraživača vjerojatno ne bi ponudila nikakve rezultate jer rangiraju samo dokumente koji sadrže sve pojmove upita. Ako se i mogu naći dokumenti koji sadrže sve riječi, suptilne distinkcije jezika koje se koriste u dugim, gramatički ispravnim upitima često će se izgubiti u rezultatima. Pretraživači koriste algoritme za rangiranje koji se temelje prvenstveno na statističkom prikazu teksta kao skupa riječi, a ne na sintaktičkim i semantičkim značajkama.⁴³

Stručnost uvelike utječe na to kako korisnici pretražuju, a razumijevanje razlika između početnika i eksperata pomaže u omogućavanju boljeg iskustva pretraživanja. Kada se govori o stručnosti korisnika, prvo treba razlučiti dvije dimenzije stručnosti. Stručnost u području domene (engl. *domain expertise*) odnosi se na poznavanje određenog predmeta, a tehnička stručnost (engl. *technical expertise*) ukazuje na znanje u korištenju računala, interneta, pretraživača i sl. Obje su dimenzije stručnosti vrijedne i korisne, ali korisnici će postići najbolje rezultate kada posjeduju obje vrste znanja. Kombinacija ove dvije dimenzije stručnosti daje podjelu koja se sastoji od četiri vrste korisnika: dvostruki eksperti, dvostruki

⁴² Sahami, M. ... et al. The happy searcher: Challenges in web information retrieval. // Pacific Rim International Conference on Artificial Intelligence / edited by Zhang, C. et. al. Berlin : Springer, 2004., str. 4.

⁴³ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 188-189.

početnici, eksperti u domeni/tehnički početnici, početnici u domeni/tehnički eksperti. Za svaku od navedenih kategorija moguće je vezati određene karakteristike.⁴⁴

Dvostruki početnici postavljaju više upita nego eksperti, ali posjete manje stranica. Iako često reformuliraju svoje upite, uglavnom se radi o malim, nedosljednim promjenama koje nemaju osobit učinak. Nakon što posjete neku stranicu, veća je vjerojatnost da će se, u odnosu na eksperte, vratiti na početnu stranicu s rezultatima zbog straha da ne „odlutaju“ predaleko. Veći broj upita i često vraćanje na stranicu s rezultatima znače da početnici na pretraživanje potroše više vremena nego što bi eksperti. Početnicima upotrebu olakšavaju korisnička sučelja koja uz upit prikazuju popis povezanih pretraživanja kako bi korisnici mogli uspješnije reformulirati svoje upite. Također, sustavi koji koriste tzv. mrvice kruha (engl. *breadcrumbs*) olakšavaju korisnicima orijentaciju. Mrvice kruha navigacijska su shema koja omogućava korisnicima da otkriju na kojem se mjestu nalaze na web stranici, istovremeno im pružajući put da se vrate na mjesto odakle su započeli pretraživanje.⁴⁵

Dvostruki eksperti ne ulaze u pretraživanje oprezno kao početnici, istražujući stranicu po stranicu, već odmah žele izravno doći na željenu destinaciju. U odnosu na početnike, oni posjete više stranica te se ne vraćaju često natrag već slijede poveznice jednu za drugom. Efikasniji su u pretraživanju i potroše manje vremena, a ne samo što preformuliraju upite rjeđe, nego i određuju relevantnost određene stranice brže. Korisničko sučelje prilagođeno ekspertima treba podržavati naprednu sintaksu i filtriranje kako bi omogućilo korisnicima da brzo suze opseg svog pretraživanja.⁴⁶

Korisnici koji su eksperti u području domene/tehnički početnici imaju širok vokabular koji im omogućava da konstruiraju upite koji su tematski relevantni. Osim toga, imaju znanje koje olakšava i ubrzava proces evaluiranja stranice. Međutim, nedostatak tehničke ekspertize uzrokuje osjećaj dezorijentiranosti među rezultatima i stranicama, sprečavajući ih da odu predaleko od početne stranice. S druge strane, korisnici koji su početnici u području domene/tehnički eksperti, vjerojatnije će upotrebljavati tehnike poput dvostrukih navodnika i Booleovih operatora kako bi oblikovali upit. Također, takvi korisnici nisu zabrinuti da bi se mogli izgubiti među rezultatima i samopouzđani su prilikom pretraživanja. Ipak, nedostaje im znanja koje omogućava određivanje relevantnosti rezultata.⁴⁷

⁴⁴ Russell-Rose, T.; Tate, T. Nav. dj., str. 4.

⁴⁵ Isto, str. 4-5.

⁴⁶ Isto, str. 6-7.

⁴⁷ Isto, str. 8-9.

8. Vrste upita na webu

Klasičan model pretraživanja informacija počiva na pretpostavci da korisnici provode pretraživanje jer imaju informacijsku potrebu. Međutim, Broder tvrdi kako razlog za pretraživanjem weba ne mora uvijek biti informacijska potreba, već da postoje i druge vrste potreba. U odnosu na potrebe koje korisnike motiviraju na pretraživanje, Broder je podijelio vrste upita u tri kategorije – informacijske, navigacijske i transakcijske.⁴⁸

Svrha je navigacijskih upita doći do određene web stranice koju korisnik ima na umu prije nego što započne pretraživanje, bilo zbog toga što ju je posjećivao u prošlosti ili zato što pretpostavlja da takva stranica postoji. Ova vrsta upita u pravilu ima samo jedan „točan“ rezultat. Informacijski upiti pronalaze informacije za koje se pretpostavlja da su dostupne na webu u statičkom obliku, te se ne predviđaju daljnje interakcije, osim čitanja. Pod statičkim oblikom misli se na to da ciljani dokument nije stvoren kao odgovor na korisnikov upit, nego je postojao otprije. Informacijski upiti mogu biti jako široki, npr. „auti“, ili uski, npr. „anemija srpastih stanica“. Kada korisnik želi doći do web mjesta na kojem će se odvijati daljnja interakcija, riječ je o transakcijskim upitima. Transakcije obuhvaćaju e-kupovinu, preuzimanje različitih vrsta datoteka (pjesme, slike, itd.), pronalaženje različitih web posredovanih usluga, pristup bazama podataka i sl. Rezultate takvih upita vrlo je teško procijeniti u smislu klasičnog pretraživanja informacija. Čimbenici koji su korisnicima bitni u ovakvim upitima, poput brzine usluge, kvaliteta slike, cijene robe i sl., uglavnom nisu dostupni pretraživačima.⁴⁹

Na ovaj su se model nadovezali Rose i Levinson, te ga djelomično izmijenili i proširili. Njihova se podjela sastoji od navigacijskih i informacijskih upita, te upita o resursima. Pomoću navigacijskog upita korisnik želi doći do početne stranice za koju je zainteresiran. Kako bi upit bio navigacijski, korisnik mora na umu imati točnu web stranicu. Iz tog razloga, većina upita koji sadržavaju imena kompanija, sveučilišta, organizacija i sl. smatraju se navigacijskima. S druge strane, potraga za osobama pomoću njihovog imena ne smatra se navigacijskim upitom. Primjerice, upisivanje imena neke poznate osobe rezultirat će u mnoštvu rezultata – stranicama fanova, člancima na portalima, itd. Malo je vjerojatno da je korisnik na umu imao specifičnu stranicu kada je proveo tu pretragu. Kada korisnik pokušava saznati više informacija o temi o kojoj pretražuje, riječ je o informacijskim upitima. Ova

⁴⁸ Broder, A. A taxonomy of web search. // ACM Sigir forum 36, 2(2002), str. 3-5.

⁴⁹ Isto, str. 5-6.

kategorija uključuje odgovore na pitanja (otvorena i zatvorena), traženje savjeta i neusmjerene upite čiji je cilj jednostavno saznati više o određenoj temi. Većina upita koji se odnose na vijesti, znanost, medicinu i povijest pripadaju kategoriji neusmjerenih upita. Informacijski upiti također uključuju lociranje nečega u stvarnom svijetu ili traženje prijedloga za buduće pretraživanje. Upiti o resursima odnose se na upite čiji je cilj dobiti nešto (osim informacija). Resurs može biti nešto što korisnik namjerava rabiti u stvarnom svijetu, poput tekstova ili akorda pjesama, recepata, šivaćih uzoraka i sl.⁵⁰

9. Društveno pretraživanje

Prije nego što je internet postao široko dostupan, traženje informacija odvijalo se u društvenom kontekstu – informacije su se dobivale od prijatelja, poznanika ili profesionalca u određenom području. Tražeći informacije na takav način možemo odrediti njihovu kvalitetu jer poznajemo kompetencije osobe od koje je informacija dobivena. Problem s mnogim današnjim web pretraživačima upravo je pitanje pouzdanosti informacija. Prema tome, kako ljudi imaju povjerenja u stvarne osobe, dok im se algoritmi koji rangiraju rezultate često čine apstraktni i nerazumljivi, čini se logičnim integrirati neku vrstu društvenog konteksta u proces pretraživanja. Prvi pokušaj uvođenja ljudske prosudbe kako bi se povećala relevantnost rezultata dogodio se 1971. kada je Rocchio uveo povratnu informaciju o relevantnosti (engl. *relevance feedback*). Riječ je o algoritmu koji omogućava korisniku da postepeno poboljšava svoj upit tako da procjenjuje relevantnost dobivenog rezultata.⁵¹ Nakon što korisnik postavi inicijalni upit, dobije određen broj dokumenata. Poslije toga odredi koji su od dobivenih dokumenata relevantni, a sustav u skladu s tim reformulira originalan upit. Zatim korisnik dobiva dokumente koji su relevantni za reformulirani upit, a proces se može nastaviti odvijati sve dok korisnikova informacijska potreba nije ostvarena.⁵² Ovo je također prva upotreba strojnog učenja u pretraživanju informacija.⁵³ Iako ovo nije nova ideja, dugo ju je vremena bilo teško učinkovito provoditi jer je potrebno navesti korisnike da dobrovoljno daju povratnu informaciju o relevantnosti rezultata, što malo korisnika želi učiniti. Različite primjene

⁵⁰ Rose, D. E.; Levinson, D. Understanding user goals in web search. // Proceedings of the 13th international conference on World Wide Web. New York : ACM, 2004., str. 14-15.

⁵¹ Burghardt, M.; Heckner, M.; Wolff, C. The many ways of searching the web together : A comparison of social search engines. // Web search engine research / edited by Dirk Lewandowski. Vol.4. United Kingdom : Emerald Group Publishing Limited, 2012., str. 20-21.

⁵² Buckley, C.; Salton, G.; Allan, J. The effect of adding relevance information in a relevance feedback environment. // SIGIR'94 / edited by Croft B.W.; van Rijsbergen C.J. London : Springer, 1994., str. 292.

⁵³ Sanderson, M.; Croft, W. B. Nav. dj., str. 1447.

društvenog pretraživanja, kao i napredak korisnika od pasivnih konzumenata do aktivnih tvoraca web sadržaja, omogućili su realizaciju ideje o povratnim informacijama i njihovo dijeljenje s drugim korisnicima. U 1990-ima počelo se razmatrati o kognitivnim i kontekstualnim aspektima tijekom procesa pretraživanja, što je označilo bitan događaj za razvoj društvenog pretraživanja. Prije toga, sustavi za pretraživanje bili su promatrani i evaluirani s teorijskog aspekta i aspekta samog sustava.⁵⁴

Najvažniji utjecaj na razvoj društvenog pretraživanja odigrali su društveni mediji i razvoj weba 2.0. Kada je web 1.0 evoluirao u web 2.0 došlo je do prijelaza iz statičnog weba u društveni. Korisnici više nisu bili samo pasivni konzumenti sadržaja, već njegovi aktivni kreatori. Dobili su novu slobodu, a omogućena im je i interakcija s ostalim korisnicima i povezivanje s osobama sličnih interesa. Najzaslužniji za ove mogućnosti, a ujedno i najkorištenije tehnologije u ovom novom razdoblju, zasigurno su društveni mediji. Oni su svojim članovima omogućili da komuniciraju jedni s drugima, izmjenjuju poruke i različit sadržaj poput slika i videa, komentiraju članke, kreiraju i ažuriraju profile, igraju online igre, itd. I ostala web mjesta, primjerice portali, postali su društveni jer su pružili opcije dijeljenja, komentiranja, i sl.⁵⁵ Pojam društveni mediji odnosi se na tehnologije, platforme i usluge koje omogućuju pojedincima da se upuste u komunikaciju jedan-na-jedan, jedan-na-mnogo, ili mnogo-na-mnogo. Najpopularniji su oblik društvenih medija društvene mreže, platforme koje okupljaju korisnike i omogućavaju im umrežavanje u virtualne zajednice na temelju zajedničkih interesa (npr. struka, kultura, hobi). Mogu biti interne (okupljati zatvoren krug ljudi) i eksterne (dostupna svim korisnicima), specijalizirane (okupljaju korisnike iz određenog područja) ili opće i okupljati sve zainteresirane korisnike (npr. Facebook).⁵⁶

Društveno pretraživanje upotrebljava kolektivnu inteligenciju korisnika weba kako bi proizvelo relevantnije rezultate pretraživanja i pomoglo korisniku da u sustavu preopterećenom dokumentima pronađe ono što traži, pritom uzimajući u obzir njegove specifične potrebe. Može ga se shvatiti i kao produžetak prirodnih obrazaca komuniciranja u stvarnom svijetu.⁵⁷ Većina tradicionalnih web pretraživača fokusira se na algoritme bazirane na sadržaju i tekstu, dok se društveni pretraživači fokusiraju na ljudsku prosudbu u rangiranju

⁵⁴ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 21.

⁵⁵ Bouadjenek, M. R.; Hacid, H.; Bouzeghoub, M. Social networks and information retrieval, how are they converging? A survey, a taxonomy and an analysis of social information retrieval approaches and platforms. // Information Systems 56, (2016), str. 1-2.

⁵⁶ Previšić, J.; Ozretić Došen, Đ. Internetski marketing. Obnovljeni separat poglavlja 19. u knjizi. Zagreb : Marketing, Adverta, 2004., str. 13.

⁵⁷ Richman, D. Nav. dj., str. 2.

i procjeni relevantnosti online izvora. To znači da algoritmi za rangiranje kod društvenih pretraživača nisu primarno bazirani na frekvenciji i distribuciji određenih ključnih riječi, već na činjenici da drugi korisnici procjenjuju određene dokumente kao zanimljive ili relevantne, obzirom na potrebu za određenim informacijama.⁵⁸ Svi korisnici weba morali bi moći pomoći u klasificiranju, organiziranju i rangiranju web stranica što ukazuje na demokratsku prirodu društvenog pretraživanja. Određeni korisnici posjeduju znanje i interes za određene teme, stoga se nameće pitanje kako organizirati povratne informacije korisnika tako da „prave osobe“ pružaju „prave odgovore“ na „prava pitanja“.⁵⁹

Principi društvenog pretraživanja primjenjuju se u različitim sustavima i aplikacijama na webu, a raznolikost primjene pokazuje kako koncept društvenog pretraživanja omogućuje širok raspon tumačenja jer mu nedostaje standardizirana i općenito prihvaćena definicija. Kako bi se društveno pretraživanje razumjelo i definiralo, potrebno je objasniti dvosmislen pojam „društveno“. Neki pretraživači su označeni kao društveni jer pretražuju društvene podatke, što se može odnositi na informacije o stvarnim osobama ili pretraživanje društvenih medija u stvarnom vremenu, dok drugi društveni pretraživači nose taj pridjev zato što koriste znanje i prosudbu ljudi pri pretraživanju weba. Pojmom društveno pretraživanje često se pokrivaju obje definicije.⁶⁰

Tako primjerice Croft, Metzler i Strohman tvrde kako društveno pretraživanje treba neku vrstu društvenog okruženja koje može biti definirano kao okruženje u kojemu zajednica korisnika aktivno sudjeluje u procesu pretraživanja.⁶¹ McDonnell i Shiri definiraju društveno pretraživanje kao upotrebu društvenih medija kao pomoć u pronalaženju informacija na internetu.⁶² Evans i sur. definiraju društveno pretraživanje kao način na koji korisnici koriste druge osobe i ostale dostupne društvene resurse tijekom procesa pretraživanja.⁶³ Vuorikari i Koper vide društveno pretraživanje kao korištenje ponašanja drugih korisnika u situacijama kada se osjećamo izgubljeno, kako bismo olakšali kretanje na webu.⁶⁴ Amitay i sur. smatraju

⁵⁸ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 20.

⁵⁹ Richman, D. Nav. dj., str. 2.

⁶⁰ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 23.

⁶¹ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 397.

⁶² McDonnell, M.; Shiri, A. Social search: A taxonomy of, and a user-centred approach to, social web search. // Program 45, 1(2011), str. 6.

⁶³ Evans, B. M.; Kairam, S.; Pirolli, P. Exploring the cognitive consequences of social search. //CHI'09 Extended Abstracts on Human Factors in Computing Systems. New York : ACM, 2009., str. 3378.

⁶⁴ Vuorikari, R.; Koper, R. Ecology of social search for learning resources. // Campus-Wide Information Systems 26, 4(2009), str. 273.

kako se društveno pretraživanje odnosi na mnoge aspekte pretraživanja weba 2.0., uključujući traganje za sličnim korisnicima na temelju sličnosti oznaka.⁶⁵

Kako bi se obuhvatile različite primjene društvenog pretraživanja, u ovom će se radu ono shvatiti kao bilo koji sustav pretraživanja informacija koji se na neki način oslanja na društveni kontekst korisnika kako bi se poboljšao proces pretraživanja.

10. Aspekti društvenog pretraživanja

Nekoliko je aspekata ključno za bolje razumijevanje termina društveno pretraživanje. Oni se odnose na načine suradnje korisnika tijekom pretraživanja, cilj pretraživanja i put kojim dolaze do rezultata.

1. Vrijeme suradnje (sinkrona i asinkrona suradnja) - Ako skupina korisnika vrši pretraživanje istovremeno, tada je pretraživanje sinkrono. Ovo omogućava da korisnici utječu jedni na druge prilikom pretraživanja, a time i na oblikovanje upita i tijek pretraživanja. Asinkronu interakciju možemo shvatiti na dva načina. U prvom obliku korisnici koriste podatke stvorene od strane drugih korisnika (npr. korisničke oznake), a u drugome skupina korisnika zajednički pretražuje, ali to ne čini u isto vrijeme. Oni koji vrše pretraživanje kasnije, mogu imati koristi od rezultata onih koji su pretraživali prije njih, dok oni koji pretražuju prvi nemaju tu mogućnost. Većina definicija društvenog pretraživanja fokusira se na asinkronu interakciju. Međutim, postoji rastući interes za istraživanje sinkrone suradnje.⁶⁶

2. Lokacija (ista lokacija i udaljena lokacija) – Kada korisnici vrše pretraživanje u paru ili grupi, mogu se nalaziti na istoj lokaciji, što im omogućuje da komuniciraju na način koji nije računalno posredovan, ili mogu biti pozicionirani na udaljenim lokacijama. Pretraživanje na istoj lokaciji vezano je uz sinkronu suradnju, dok pretraživanje na udaljenim lokacijama može biti sinkrono i asinkrono. Ako se korisnici nalaze na različitim lokacijama, komuniciraju putem poruka, e-pošte, poziva i sl.⁶⁷

3. Oblik suradnje (implicitna i eksplicitna suradnja) – Društvena je suradnja implicitna u sustavima u kojima proizlazi iz analize ponašanja mnogih korisnika, ali sami korisnici nikada ne dolaze u interakciju jedni s drugima. Ona nastaje kada ljudi konzumiraju ono čemu su drugi prethodno pridonijeli. Primjerice, neki sustavi upotrebljavaju podatke o

⁶⁵ Amitay, E. ... et. al. Social search and discovery using a unified approach. // Proceedings of the 20th ACM conference on hypertext and hypermedia. New York : ACM, 2009., str. 199.

⁶⁶ McDonnell, M.; Shiri, A. Nav. dj., str. 9.

⁶⁷ Isto, str. 10.

korisničkim oznakama kako bi prilagodili rangiranje web pretraživanja. Takav sustav uvelike ovisi o društvenom produktu mnogih korisnika i stoga se može smatrati suradničkim. Međutim, on sam ne pomaže suradnji korisnika, nego jednostavno iskorištava podatke stvorene iz drugih procesa suradnje. U eksplicitnim suradnjama ljudi aktivno dijele svoja mišljenja, uključujući označavanje ili klasifikaciju web stranica radi koristi drugih korisnika.⁶⁸

4. Cilj pretraživanja (pronalaženje osoba i pronalaženje informacija) – Neki istraživači definiraju društveno pretraživanje kao potragu za osobama. Primjerice, postoje stranice koje omogućavaju potragu za osobama na temelju njihovih karakteristika poput interesa, društvenih krugova, institucija, itd., ali društveno se pretraživanje rjeđe fokusira na taj aspekt. Većina se perspektiva usredotočuje na pronalaženje informacija kao primarni cilj korisnika. Društvene interakcije koje se pritom odvijaju samo su sredstvo dolaženja do cilja. Čak i oni koji definiraju društveno pretraživanje kao potragu za osobama, često naglašavaju kako je krajnji cilj pronalazak informacija.⁶⁹

5. Dolazak do rezultata (pretraživanje i pronalazak) – Pretraživanje se odnosi na tradicionalni model pretraživanja informacija u kojemu se smatra kako korisnik zna što traži i sposoban je postaviti odgovarajući upit kako bi dobio željene rezultate. S druge strane, do otkrića dolazi kada naiđemo na nešto što nismo ni tražili. Na primjer, korisnici mogu pregledavanjem oznaka drugih korisnika na nekoj od stranica za korisničko označavanje pronaći vrijedan sadržaj kojeg inače ne bi pronašli.⁷⁰

11. Implementacija društvenog pretraživanja

11.1. Korisničko označavanje

Servisi za korisničko označavanje nude opcije spremanja (engl. *bookmarking*) i označivanja (engl. *tagging*) online izvora, te tako omogućuju korisnicima da sačuvaju informacije privatno, ali i da ih podijele javno s drugima. Neke od najpoznatijih stranica za korisničko označavanje su Del.icio.us i Flickr, a s radom su započele početkom 2000-ih. Servisi za korisničko označavanje besplatni su, a nakon što se korisnici registriraju, mogu odmah započeti s označivanjem, dijeljenjem i spremanjem web stranica. Korisnici čak mogu bilježiti stranice onako kako to katalogizatori rade u knjižnicama. Ipak, u praksi većina

⁶⁸ Isto, str. 11.

⁶⁹ Isto, str. 12.

⁷⁰ Isto, str. 13.

korisnika samo označava stranice s nekoliko oznaka koje se odnose na njihov sadržaj.⁷¹ Svaka oznaka predstavlja pojam koji opisuje određenu stavku (primjerice, slika cvijeća može biti označena oznakama „cvijet“, „trajnica“, „biljka“ i sl.). Na ovaj će način neki predmeti dobiti oznake, a neki će ostati neoznačeni, jer korisnici najvjerojatnije neće označiti sav postojeći sadržaj, već onaj koji ih zanima.⁷²

Korisničko označivanje obično nije uopće kontrolirano ili je slabo kontrolirano, a korisnik pojmova ne bira iz kontroliranog rječnika već iz prirodnog jezika. S obzirom na nedostatak kontrole, korisničke oznake su šarolike i može ih se podijeliti u nekoliko kategorija:⁷³

1. Oznake na temelju sadržaja – opisuju sadržaj predmeta (npr. „mačka“, „most“, „crkva“)

2. Oznake na temelju konteksta – opisuju kontekst predmeta (npr. „Dioklecijanova palača“, „Split“)

3. Oznake atributa – oznake koje opisuju implicitne karakteristike predmeta (npr. „Sony“ – vrsta kamere, „crno i bijelo“ – način fotografije)

4. Subjektivne oznake – oznake koje subjektivno opisuju predmet (npr. „fino“, „odlično“, „prekrasno“)

5. Organizacijske oznake – oznake koje pomažu korisniku u organiziranju predmeta (npr. „pročitano“, „moje slike“)

Stranice za korisničko označivanje važne su zato što su korisnici zaduženi za definiranje onoga što im se sviđa. Njihove oznake pomažu u stvaranju folksonomija, korisnički generiranih taksonomija koje pomažu u budućem pretraživanju. Stoga, vrijedna stranica koja je loše rangirana u klasičnom pretraživanju, može postati lakše dostupna jer ju je identificirala i podijelila grupa korisnika.⁷⁴ Sljedeća je prednost korisničkog označivanja ta što oznake pružaju tekstualnu dimenziju ne-tekstualnom sadržaju koji je inače teško pretraživ, poput slika i videa.⁷⁵ Osim u pretraživanju, oznake mogu biti korisne i u pregledavanju (engl. *browsing*) gdje omogućuju korisnicima da otkriju i istraže nove predmete u mnoštvu web sadržaja. Kada se korisnik nalazi na web stranici i gleda određeni predmet, sve njegove

⁷¹ Richman, D. Nav. dj., str. 5.

⁷² Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 400.

⁷³ Isto, str. 401-402.

⁷⁴ Richman, D. Nav. dj., str. 5.

⁷⁵ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 402.

oznake mogu biti prikazane, a korisnik može klikom na željenu oznaku otkrivati nov sadržaj. Pregledavanje na ovaj način donekle je ograničeno jer korisnik slijedi oznake koje su sadržajno povezane što mu omogućava da otkrije samo određeni skup predmeta. Primjerice, ako je krenuo od slike psa, to će ga teško dovesti od turističkih vodiča. Jedan od načina na koji korisnik može dobiti širi pregled oznaka je pomoću najpopularnijih oznaka. One se mogu odnositi na cijelu stranicu ili na određenu kategoriju predmeta.⁷⁶

Nedostatak korisničkog označivanja je taj da ono zahtijeva određeno razumijevanje i učenje kako bi ga se upotrebljavalo. Samo mali broj korisnika odvoji vrijeme kako bi označio sadržaj pa zato neka pretraživanja mogu reflektirati interese i sklonosti manje skupine korisnika. Također, neki korisnici spamaju dodajući netočne i nepovezane oznake sadržaju. Takav oblik ponašanja teško je kontrolirati.⁷⁷ Kako bi se izbjegao problem spamova, neprikladnih ili pogrešnih oznaka, potrebno je potaknuti korisnike da unose kvalitetne oznake. Jedan od načina kojim bi se ovo moglo postići je dozvoliti korisnicima da prijave neprikladne oznake. Osim toga, sustav im može omogućiti neke privilegije ili poboljšani status ako unesu određen broj kvalitetnih oznaka kroz određeni period.⁷⁸

Za razliku od korisničkog označivanja, pretraživači koriste automatsko indeksiranje kako bi dodijelili identifikatore dokumentima na webu.⁷⁹ Veliki problem indeksiranja je taj što su indekseri u većini slučajeva potpuno odvojeni od procesa pretraživanja, odnosno profesionalni indekser ili autor dokumenta obično nije taj koji će tražiti navedeni dokument. To može otežati proces pretraživanja. Kada korisnik treba formulirati upit koji se bazira na informacijama koje želi pronaći, dolazi do razilaženje u vokabularu između indeksera i korisnika.⁸⁰ Prednosti automatskog indeksiranja su da je ono iscrpno, što znači da je svaki pojam u dokumentu indeksiran i ništa nije izostavljeno, te da je konzistentno, za razliku od korisničkog označivanje koje može biti pristrano i sklono pogreškama. Istraživanja u području indeksiranja pokazuju da je automatsko indeksiranje barem jednako učinkovito kao manualno, a često i mnogo bolje. Međutim, rezultati pokazuju i da se ova dva oblika indeksiranja međusobno nadopunjuju i da će pretraživanje biti najučinkovitije uz korištenje oba oblika.⁸¹

⁷⁶ Isto, str. 406.

⁷⁷ Richman, D. Nav. dj., str. 6.

⁷⁸ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 404.

⁷⁹ Isto, str. 400.

⁸⁰ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 26.

⁸¹ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 400.

11.2. Sustavi pitanja i odgovora

Sustavi pitanja i odgovora (engl. *question answering systems*) predstavljaju posebnu vrstu pretraživača, koji se od ostalih sustava za pronalaženje informacija najviše razlikuje u načinu na koji korisnici formuliraju svoje informacijske upite. Korisnici pitanje postavljaju prirodnim jezikom, što znači da nisu prisiljeni prevoditi upit u manje intuitivne deskriptore i operatore. Dobiveni rezultati također su izraženi prirodnim jezikom i neposredan su odgovor na korisnikov upit, a ne popis relevantnih ili zanimljivih web stranica. Navedene su značajke ujedno i najveća prednost sustava pitanja i odgovora.⁸²

Ovi se sustavi ponašaju kao posrednici između korisnika koji postavlja upit i korisnika koji bi mogao znati odgovor na to pitanje. Također pružaju mogućnost za interakciju, procjenu i rangiranje pitanja i odgovora. Na ovaj način korisnik ne dobiva samo odgovore od stvarnih osoba, već ima priliku stupiti u kontakt s istomišljenicima ili stručnjacima u određenom području. Još jedna ključna značajka sustava pitanja i odgovora je mogućnost postavljanja čak i složenih pitanja izravno, bez potrebe da se prevode u obradiva potpitanja koja bi stroj mogao obraditi.⁸³

Većinu sustava pitanja i odgovora karakteriziraju tri dimenzije: vrijeme, trošak i društveno okruženje. U odnosu na vrijeme, komunikacija može biti sinkrona, odnosno korisnik koji postavlja pitanje i onaj koji odgovara nalaze se online u isto vrijeme, što znači da će postavljatelj pitanja svoj odgovor dobiti odmah. Ako korisnik odgovor na pitanje dobije s određenim vremenskim odmakom, tada je riječ o asinkronoj komunikaciji. Postavljanje pitanja i dobivanje odgovora može biti besplatno, ali se može i naplaćivati, ovisno o sustavu koji pruža uslugu te o tome tko su korisnici koji daju odgovore na pitanja. Kada je riječ o društvenom okruženju, moguće je razlikovati zajednicu korisnika i zajednicu eksperata kao korisnike koji pružaju odgovore. Sustavi koji se temelje na zajednici korisnika obično funkcioniraju tako da omogućuje registriranim korisnicima postavljanje pitanja i odgovaranje na pitanja drugih korisnika. Korisnici mogu pretraživati odgovorena i neodgovorena pitanja putem tražilice ili pregledavati različite kategorije pitanja. Također mogu ocjenjivati i komentirati pitanja i postojeće odgovore. Neki sustavi, poput Wiki Answers, ne samo što omogućuju komentiranje danih odgovora, već dopuštaju editiranje i ispravljanje odgovora. Ovakva sloboda može dovesti do pojave vandalizma ili sukoba oko editiranja. Kako bi se te pojave izbjegle, sustavi često pokušavaju interno motivirati korisnike putem nagrađivanja

⁸² Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 29.

⁸³ Isto, str. 30.

bodovima. Više bodova označava bolju poziciju koja donosi dodatne mogućnosti, poput postavljanja više pitanje ili komentiranja na više postavljenih pitanja.⁸⁴

S druge strane, sustavi koji se temelje na zajednici eksperata obično su dostupni uz novčanu naknadu, a na pitanja odgovara zatvorena zajednica eksperata koja je stručna u određenom području. Eksperti mogu biti unaprijed odabrani od strane operatera sustava, ili se mogu sami prijaviti u sustav kao stručnjaci. Kvaliteta pitanja i odgovora ovdje je obično viša nego u otvorenim sustavima.⁸⁵

Prednosti sustava pitanja i odgovora su mogućnost dobivanja odgovora na kompleksnu informacijsku potrebu, dobivanje više različitih mišljenja, interakcija s korisnicima koji dijele interese, probleme i ciljeve.⁸⁶ Osnovni problem ovih sustava je taj što ne mogu jamčiti da je dobiveni odgovor točan. Mehanizmi poput ocjenjivanja korisnika koji su ponudili odgovor na pitanje, kao i pozicija korisnika koji je odgovorio na pitanje u društvenom krugu korisnika koji postavlja pitanje, pomažu procijeniti kvalitetu i točnost odgovora. Mana asinkronih sustava je ta što korisnik ne zna kada će dobiti odgovor na pitanje te hoće li ga uopće dobiti. Problem otvorenih sustava predstavlja loša kvaliteta odgovora koja je povezana sa samom kvalitetom postavljenih pitanja. Mogućnost postavljanja i odgovaranja na pitanja prirodnim jezikom nije uvijek djelotvorna, već potiče neformalnu, *off-topic* interakciju, kojoj nedostaje neutralno stajalište i koja podsjeća na interakciju na forumima.⁸⁷

Jedan od poznatijih sustava pitanja i odgovora bio je Aardvark. Postojao je od 2007., da bi ga Google kupio 2010., a ugasio 2011. Njegov cilj bio je povezati korisnike koji su postavili pitanje s korisnicima koji su im prijatelji na društvenim mrežama, ili prijateljima prijatelja, koji najvjerojatnije znaju odgovor na postavljeno pitanje. Kada se novi korisnik počeo koristiti Aardvarkom, sustav je provodio niz koraka indeksiranja kako bi uputio odgovarajuća pitanja odgovarajućem korisniku. Pitanja u Aardvarku preusmjeravala su se do korisnikove mreže, stoga je prvi korak bio indeksiranje prijatelja, poznanika i informacija o pripadnosti, a to se ostvarivalo pomoću društvenog grafikona (engl. *social graph*). Društveni je grafikon model društvene mreže, odnosno njime su predstavljeni društveni odnosi između korisnika. Cilj Aardvarka nije bio izgraditi društvenu mrežu, već omogućiti ljudima da koriste svoje postojeće mreže odnosa. Kako bi se to ostvarilo, korisnik je u procesu registracije mogao uvesti svoje kontakte iz društvenih mreža poput Facebooka ili LinkedIna, ili

⁸⁴ Isto, str. 30-32.

⁸⁵ Isto, str. 32-33.

⁸⁶ Croft, W. B.; Metzler, D.; Strohman, T. Nav. dj., str. 415.

⁸⁷ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 30.

pojedinačno pozvati prijatelje da se pridruže. Također, korisnici Aardvarka bili su povezani putem zajedničkih grupa koje odražavaju njihove pripadnosti u stvarnom svijetu, kao što su škole koje su pohađali ili radna mjesta. Osim toga, Aardvark je indeksirao teme o kojima korisnik posjeduje određeno znanje što ostvarivalo na više načina. Primjerice, korisnik je mogao označiti teme za koje smatra da je stručan, ili je sam sustav promatrao korisnikovo ponašanje i utvrđivao njegovo znanje. Što su korisnici bili aktivniji, to je postojalo više potencijalnih odgovora, a što su korisničke mreže bile gušće, tim je veća bila baza znanja. Kako bi to bilo ostvarivo, strategija Aardvarka bila je stvoriti dobro iskustvo za korisnike kako bi ostali aktivni i pozivali prijatelje.⁸⁸

Aardvark je sadržavao četiri glavne komponente: web pauke ili robote i indeksere, analizator upita, rangiranje i korisničko sučelje. Web roboti i indekseri pronalazili su i označavali resurse koji sadrže podatke, u ovom slučaju korisnike, a ne dokumente kao u tradicionalnom pretraživanju. U ovom se koraku ogleda najteži izazov Aardvarka - odrediti pravog korisnika koji će odgovoriti na postavljeno pitanje. Kako bi sustav bio uspješan u tome, morao je znati dvije ključne stvari o svojim korisnicima: 1) teme o kojima korisnik posjeduje znanje; 2) drugi korisnici s kojima je osoba povezana. Aardvark je određivao teme u kojima je korisnik stručan iz više izvora. Kao prvo, na samom se početku od korisnika tražilo da navedu barem tri teme o kojima posjeduju znanje. Od korisnikovih se prijatelja tražilo da navedu područja u kojima vjeruju mišljenju toga korisnika. Analizirali su se i korisnikovi profili na društvenim mrežama i blogovi (ako ih je imao), te je algoritam izdvajao teme koje se ondje spominju. Također, sustav je automatski izdvajao teme iz poruka koje korisnici šalju jedni drugima. S vremenom, Aardvark je učio koje teme ne treba povezivati s kojim korisnicima. To se ostvarivalo na tri načina: 1) kada je korisnik eksplicitno „isključio“ neku temu, 2) kada je korisnik odbio odgovoriti na postavljeno pitanje, 3) kada je korisnik dobio negativnu ocjenu svog odgovora. Aardvark je, osim kroz osobne odnose na društvenim mrežama, izračunavao povezanost korisnika na više načina: putem demografske sličnosti, sličnosti profila (npr. isti ukus u filmovima, glazbi i sl.), sličnosti vokabulara (npr. upotreba kratica), sličnosti u načinu chatanja, dužine poruka, razine ljubaznosti, brzine odgovaranja na poruke.⁸⁹

Sljedeća je komponenta sustava bila analizator upita koji je služio kako bi se shvatila informacijska potreba korisnika. Kada je korisnik postavio pitanje, prvo se na njega

⁸⁸ Horowitz, D.; Kamvar, S. D. The anatomy of a large-scale social search engine. // Proceedings of the 19th international conference on World wide web / edited by Rappa, M. New York : ACM, 2010., str. 432-433.

⁸⁹ Isto, str. 433-434.

primjenjivalo nekoliko klasifikatora koji su određivali da li je pitanje uistinu pitanje ili je krivo poslana poruka, slijed ključnih riječi i sl, a u tom se slučaju korisnika tražilo da ponovi pitanje. Zatim se provjeravalo da li je sadržaj upita primjeren te nije li poruka spam, a ako je o tome riječ korisnika se upozoravalo te također tražilo da ponovi pitanje. Ako je pitanje bilo jednostavan faktografski upit, korisniku se nudio odgovor koji proizlazi iz tradicionalne internetske pretrage. Gledalo se i da li je pitanje geografski određeno (npr. „Koji je najbolji kineski restoran u Zagrebu?“) kako bi na pitanje mogao odgovoriti korisnik koji je povezan s tom lokacijom.⁹⁰

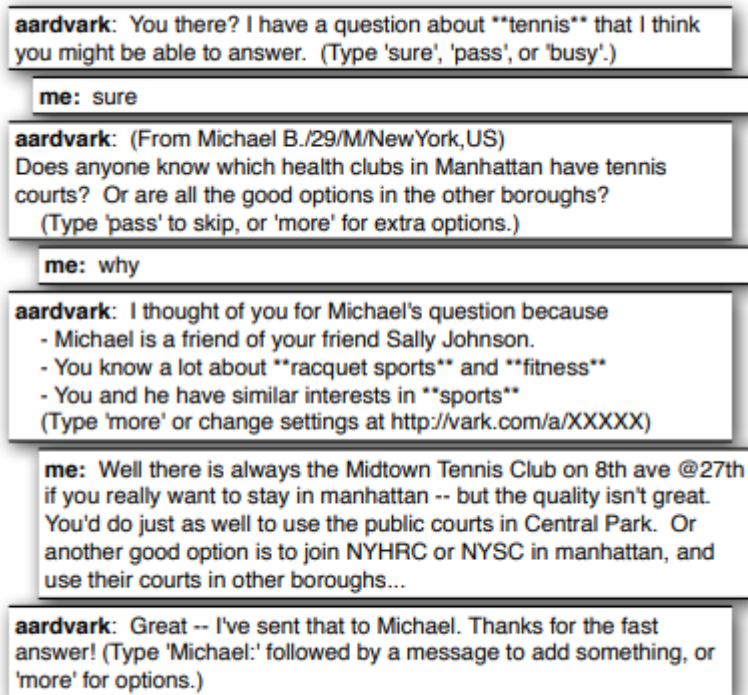
Algoritam za rangiranje u Aardvarku utvrđivao je koji su korisnici pogodni kandidati za odgovaranje na određeno pitanje, uzimajući u obzir tko je osoba koja je postavila pitanje, sadržaj pitanja te dostupnost korisnika. Kako bi korisnik što prije dobio odgovor na postavljeno pitanje, sustav je davao prednost korisnicima koji su u tom trenutku bili online, koji su obično u to doba dana bili online ili koje sustav već neko vrijeme nije kontaktirao sa zahtjevom da odgovore na pitanje.⁹¹

Korisničko je sučelje bilo dizajnirano po uzoru na postojeće popularne komunikacijske kanale poput instant poruka, SMS poruka, e-pošte i sl. Sastojalo se od ulaznih mehanizama gdje je korisnik unosio tekst svog pitanja, te odgovora kojeg korisnik dobije od sustava. Sustav je pitanje upućivao korisniku koji je mogao potvrditi da želi odgovoriti na pitanje ili odbiti., što je vidljivo na slici 2. Ako je odgovorio potvrdno, sustav mu je prosljeđivao pitanje, a korisnik je mogao zatražiti i dodatna pojašnjenja vezana uz pitanje. Sustav je nakon toga njegov odgovor prosljeđivao postavljaju pitanja. Korisnicima je bila osigurana privatnost, nitko drugi nije mogao vidjeti njihove odgovore na pitanja osim osobe koja je pitanje postavila, zbog čega se nisu morali ustručavati u davanju odgovora, mogli su biti iskreni i slobodni. Korisnik je na dobiveno pitanje mogao odgovoriti i „zašto“ („*why*“), nakon čega mu je sustav objasnio prema kojim je kriterijima upravo on izabran za to pitanje. Ako korisnik nije želio više dobivati ovakva pitanja u budućnosti, lako je mogao promijeniti postavke svog profila. Posebno se pazilo da korisnicima zahtjevi za odgovaranje na pitanja ne pristižu prečesto, kako ih sustav ne bi opteretio, a korisnik je mogao u postavkama profila odrediti koliko često želi dobivati pitanja.⁹²

⁹⁰ Isto, str. 434.

⁹¹ Isto, str. 435.

⁹² Isto, str. 435-436.



Slika 2. Primjer interakcije Aardvarka i korisnika⁹³

11. 3. Zajedničko pretraživanje

U zajedničkom pretraživanju (engl. *collaborative search*) sudjeluju korisnici koji dijele informacijsku potrebu i aktivno rade zajedno kako bi ju ispunili. Sudionici pretraživanja mogu s ostalim korisnicima dijeliti svoja iskustva pretraživanja, tj. svoje upite i rezultate, kao bi ostalima olakšali pretraživanje i pomogli im da preformuliraju upite. Korisnici koji zajedno pretražuju uglavnom se poznaju otprije, a ovaj je oblik pretraživanja najpogodniji za kompleksna pretraživanja čiji dijelovi mogu biti podijeljeni između pojedinaca. Stoga je ključna stavka sustava zajedničkog pretraživanja mehanizam koji će koordinirati kombinirana pretraživanja različitih korisnika. Korisnici moraju imati vizualan prikaz pretraživanja drugih korisnika, potrebno im je omogućiti prikaz relevantnih pretraživanja ostalih korisnika s kojima surađuju kako bi se izbjegla nepotrebna pretraživanja. Osim toga, kako bi proces pretraživanja bio zadovoljavajuć za korisnike, moraju se moći pridružiti usred pretraživanja ili se okrenuti nekim drugim smjerovima u zadanom smjeru pretraživanja. Na zajedničko se pretraživanje može gledati kao na vrstu društvenog pretraživanja gdje je suradnja

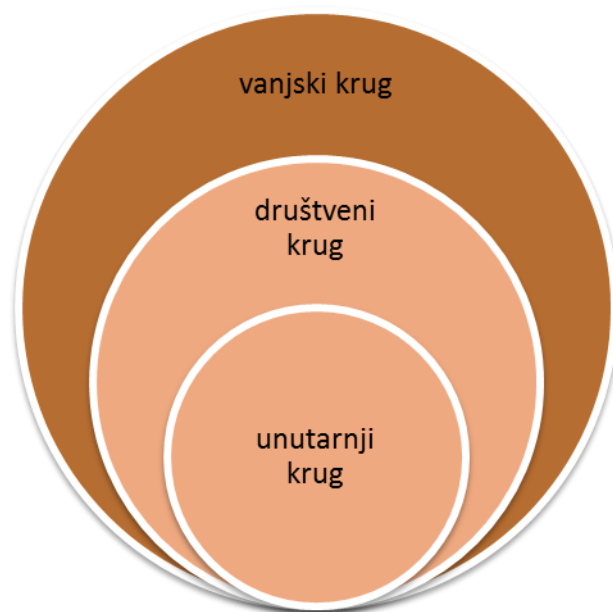
⁹³ Isto, str. 436.

najeksplicitnija. Kao podskupinu zajedničkog pretraživanja možemo smatrati zajedničko pregledavanje (engl. *collaborative browsing*). Ono može poprimiti različite oblike – traženje određene informacije, istraživanje prethodno neistraženog sadržaja kako bi otkrili nešto zanimljivo ili kombinacija navedenog.⁹⁴

Najveći problem zajedničkog pretraživanja jest apsolutna transparentnost postupaka svakog korisnika tijekom procesa suradnje. Naime, upiti svih dostupni su svima, stoga korisnici mogu biti nepovjerljivi i ustručavati se postaviti upit za koji smatraju da bi mogao biti glup, suvišan ili naivan.⁹⁵

Pretraživanje se može vršiti na zajedničkoj lokaciji, ili na različitim lokacijama. Ako se ono odvija na različitim lokacijama, tada postoji potreba za dodatnim kanalima putem kojih će se koordinirati aktivnosti korisnika, poput chata, video ili audio konferencije.⁹⁶

Zajedničko pretraživanje može se promatrati kroz nekoliko razina suradnje, što je prikazano na slici 3.



Slika 3. Razine suradnje u zajedničkom pretraživanju⁹⁷

⁹⁴ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 34.

⁹⁵ Isto, str. 35.

⁹⁶ Golovchinsky, G.; Pickens, J.; Back, M. A taxonomy of collaboration in online information seeking. // Search 32, (2009), str. 2.

⁹⁷ Russell-Rose, T.; Tate, T. Nav. dj., str. 254.

U sredini se nalazi unutarnji krug koji predstavlja jezgru suradnje, a čine ga korisnici koji dijele zadatak i imaju zajednički cilj. Informacijska potreba svih sudionika ne mora biti identična, ovisit će o tome kako su raspodijelili dužnosti, a moguće je da i samo jedan korisnik vrši pretraživanje dok ostali promatraju i doprinose svojim idejama i savjetima. Svaki sudionik pretraživanja u unutarnjem krugu vjerojatno je povezan s raznim društvenim grupama. One mogu biti povezane zajedničkim interesima, geografskom lokacijom, poslom i sl. Tijekom procesa pretraživanja korisnici se mogu obratiti pojedincima iz jednog ili više društvenih krugova kojima pripadaju kako bi ih zamolili za pomoć.⁹⁸ Kada tražimo informacije od drugih osoba, možemo imati različite koristi. Prva i očita korist je da dobijemo odgovor na pitanje ili rješenje za problem. Znanje koje se pritom prenosi može biti deklarativno, odnosno opisi informacija koje su neophodne za rješavanje problema i veza između njih, ili proceduralno koje se odnosi na procedure i postupke za uporabu znanja. Ako i ne dobijemo odgovor na naše pitanje, možemo steći meta-znanje. Osoba od koje tražimo informaciju može nas uputiti na drugi izvor koji posjeduje tu informaciju, bilo da je riječ o osobi ili bazi podataka. Ako znamo lokaciju informacije, time se povećava mogućnost rješavanja problema i pospješuje efikasnost. Od pomoći se može pokazati i razgovor s drugima o problemu jer nam može pružiti širu perspektivu, otvoriti nove dimenzije i utjecati na to da preformuliramo naš problem. Sama interakcija može biti vrijedna i kada ne dobijemo odgovor, jer smo od sugovornika dobili potvrdu da su naši planovi i ideje dobri, što jača samopouzdanje i pomaže da sami sebi opravdamo svoje odluke. Naposljetku, sam autoritet osobe od koje smo dobili informaciju nosi određeno značenje. Ako možemo reći da smo konzultirali osobu čije je stručnost u određenom području cijenjena, to će povećati kredibilitet same informacije.⁹⁹

Unutarnji krug i društveni krug odražavaju eksplicitni, namjerni oblik suradnje između pojedinaca koji su barem u nekoj mjeri društveno povezani. S druge strane, vanjski je krug područje implicitne, nenamjerne suradnje među pretežno strancima. U ovom slučaju pretraživači upotrebljavaju veliku količinu korisnički generiranih podataka kako bi korisnicima pružili relevantne rezultate. Pretraživači mogu koristiti tehnike koje su u ovoj podjeli obično povezane s vanjskim krugom (suradničko filtriranje, analiza klikova, itd.) i primijeniti ih u društvenom krugu korisnika kako bi pružili personalizirane rezultate

⁹⁸ Isto, str. 254-255.

⁹⁹ Cross, R.; Rice, R. E.; Parker, A. Information seeking in social context: Structural influences and receipt of information benefits. // IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews) 31, 4(2001), str. 439-440.

pretraživanja. Primjerice, Google je upotrebljavao podatke s društvene mreže Google+ kako bi poboljšao rezultate.¹⁰⁰

Jedan je od primjera zajedničkog pretraživanja sustav CoSearch kojeg su razvili Amershi i Morris 2008., a služi za zajedničko pretraživanje na istoj lokaciji. Sustav se sastoji od primarnog zaslona računala, tipkovnice i miša, a dodatna oprema uključuje pametne telefone i višestruke miševe. Osoba koja je vođa tima i kontrolira pretraživanje naziva se upravljač (engl. *driver*), dok su ostali sudionici promatrači (engl. *observers*). Proces pretraživanja započinje kada upravljač postavi upit u tražilicu, a rezultati pretraživanja prikazuju se na primarnom zaslonu i na mobilnim telefonima promatrača. Oni mogu kliknuti na rezultate pretraživanja čime će se odabrana stranica pohraniti u red čekanja (engl. *page queue*), što omogućava svakom sudioniku da sugerira koju bi stranicu trebalo posjetiti sljedeću. Osim reda čekanja, postoji i red upita (engl. *query queue*) gdje sudionici pohranjuju nove upite koje će istražiti u nastavku pretraživanja.¹⁰¹

Sljedeći je primjer sustav SearchTogether u kojemu je riječ o zajedničkom pretraživanju na udaljenim lokacijama. Sustav su 2007. razvili Morris i Horovitz, a u njemu se pretpostavlja da se sudionici nalaze na različitim lokacijama te da svaki od njih posjeduje svoje računalo. Za razliku od zajedničkog pretraživanja na istoj lokaciji gdje su svi sudionici prisutni dok se odvija proces pretraživanja, kod pretraživanja na udaljenim lokacijama sudionici ne moraju biti online u isto vrijeme. Sustav funkcionira tako da sudionici postavljaju upite koji su vidljivi svim sudionicima što omogućava ostalima da vide kako teče proces pretraživanja kod drugih nakon čega mogu promijeniti ili preformulirati svoje upite. Sudionici mogu ocjenjivati i komentirati stranice koje su drugi posjetili, a sudionik može drugom sudioniku preporučiti stranicu za koju smatra da bi ju ovaj trebao posjetiti. U slučaju da se novi sudionik pridruži pretraživanju, lako može pregledati prethodne upite, ocijenjene stranice, komentare i preporuke te tako biti u toku.¹⁰²

11.4. Suradničko filtriranje i sustavi preporuke

Suradničko filtriranje (engl. *collaborative filtering*) metoda je filtriranja koja koristi informacije o vezama između korisnika, ili između njihovih profila, kako bi automatski

¹⁰⁰ Russell-Rose, T.; Tate, T. Nav. dj., str. 256-257.

¹⁰¹ Amershi, S.; Morris, M. R. CoSearch: a system for co-located collaborative web search. // Proceedings of the SIGCHI conference on human factors in computing systems. New York : ACM, 2008., str. 1649 – 1651.

¹⁰² Morris, M. R.; Horvitz, E. SearchTogether: an interface for collaborative web search. // Proceedings of the 20th annual ACM symposium on User interface software and technology. New York : ACM, 2007., str. 6-8.

predvidjela interese korisnika. Pretpostavlja se da će između korisnika koji imaju isto mišljenje ili ukus u određenom području, postojati sličnosti i u drugim aspektima. Suradničko se filtriranje obično rabi za implementaciju sustava preporuke (engl. *recommender systems*).¹⁰³ Sustavi preporuke programski su alati i tehnike čiji je cilj predložiti korisniku sadržaj ili artikl za kojeg bi mogao biti zainteresiran, a primarno su usmjereni prema korisnicima kojima nedostaje osobnog iskustva ili sposobnosti da procijene i izaberu između velike količine sadržaja koju internet pruža.¹⁰⁴ Oni korisniku sugeriraju potencijalno relevantan sadržaj koji se ne bi pronašao putem korisnikovog izvornog upita. Relevantan se sadržaj selektira automatski koristeći prosudbu sličnih korisnika.¹⁰⁵

Suradničko filtriranje funkcionira tako da korisnici ocjenjuju određene stavke (knjige, filmove, serije i sl.) kako bi izrazili svoj interes, a te ocjene predstavljaju približan prikaz korisnikovog ukusa u određenom području. Nakon toga sustav uspoređuje ocjene s ocjenama ostalih korisnika kako bi sastavio listu korisnika sa sličnim ukusom, a zatim korisniku preporuči predmete koje su njemu najbliži korisnici visoko ocijenili, a on sam ih još nije ocijenio. Korisnik odmah može označiti predložene predmete koji ga ne zanimaju čime sustav nastavlja učiti o njegovim preferencijama i razvijati njegov profil. Kako bi sustav mogao povezati korisnike sličnih interesa i na osnovu toga stvoriti preporuku, moraju biti ostvarena tri uvjeta: 1) mora sudjelovati mnogo korisnika kako bi bilo vjerojatno da će svatko pronaći nekoga sa sličnim interesima, 2) sustav mora omogućiti jednostavan način na koji će korisnici izraziti svoje interese, 3) algoritmi moraju moći povezati korisnike sa sličnim interesima.¹⁰⁶

Glavna je prednost sustava preporuke to što su preporuke personalizirane, odnosno korisnik će preporučene stavke vjerojatno smatrati zanimljivima. Osim toga, sam korisnik ne mora tražiti preporuku ili osobu koja će mu nešto preporučiti, već je dovoljno da ocijeni određen broj stavki kako bi sustav odredio njegov ukus. Sustavi preporuke ne uključuju kontakt između korisnika, osim ako samim dizajnom sustava nije drukčije određeno, odnosno osim ako ne postoji opcija povezivanja i interakcije s najbližim „susjedima“. Jedan od problema sustava preporuke je problem prvog ocjenjivača (engl. *first rater*). Određena stavka ne može biti preporučena sve dok je netko nije ocijenio. Nadalje, ako je broj korisnika koji su

¹⁰³ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 36.

¹⁰⁴ Ricci, F.; Rokach, L.; Shapira, B. Introduction to recommender systems handbook. // Recommender systems handbook / edited by Rokach, L. et. al. Boston : Springer, 2011., str. 1.

¹⁰⁵ Burghardt, M.; Heckner, M.; Wolff, C. Nav. dj., str. 36.

¹⁰⁶ Terveen, L.; Hill, W. Beyond recommender systems: Helping people help each other. // Human-Computer Interaction in the New Millennium / edited by Carroll, J.M. 1st ed. New York : ACM Press, 2001., str. 12.

ocijenili stavke relativno malen u odnosu na broj stavki u bazi, vjerojatno neće postojati značajno podudaranje između korisnika, a dobivene preporuke neće biti zadovoljavajuće. Posljednji je problem taj što sustav korisniku često daje informacije koje za njega nisu nove. Primjerice, nekome tko je ocijenio jednu Shakespeareovu dramu ponudit će sve ostale, iako je vjerojatno da je korisnik koji je upoznat s jednim Shakespeareovim djelom upoznat i s ostalima. Jedno od rješenja kojim bi se sustav mogao poboljšati je izrada algoritma koji bi mogao predvidjeti ove korelacije.¹⁰⁷

Mnoge velike komercijalne web stranice kao što su Amazon ili Netflix upotrebljavaju sustave preporuke s ciljem da korisnici pronađu neki novi proizvod koji će ih zanimati i kupe ga. Stoga su sustavi preporuke korisni kompanijama, kao i krajnjim korisnicima. Pružatelji usluga mogu korištenjem sustava preporuke ostvariti različite ciljeve:¹⁰⁸

1. Povećati broj prodanih predmeta – Vjerojatno najvažnija funkcija upotrebe komercijalnih sustava preporuke je prodati veću količinu predmeta u usporedbi s količinom prodanom bez preporuke. Nekomercijalni sustavi imaju slične ciljeve, iako korisnici na njihovim stranicama ne troše svoj novac. Primjerice, cilj je portala s vijestima povećati broj pročitanih članaka na svojoj stranici.

2. Prodati više različitih predmeta – Sustav preporuke omogućava korisniku da pronađe sadržaj kojeg bi teško pronašao samostalno. Na primjer, cilj je Netflix iznajmiti sve svoje filmove i serije, a ne samo najpopularnije. Preporučivanje pogrešnih naslova pogrešnom pretplatniku nosi rizik, jer u slučaju nezadovoljstva može otkazati svoju pretplatu. Sustav preporuke osigurava da manje popularni naslovi budu preporučeni odgovarajućim pretplatnicima.

3. Povećati zadovoljstvo korisnika – Pomoću preporuke korisnici mogu pronaći relevantan sadržaj koji će zadovoljiti njihove korisničke potrebe i interes, a time će se povećati korištenje određene stranice ili servisa.

4. Povećati vjernost korisnika – Cilj je svakog web mjesta zadržati svoje posjetitelje, a kada govorimo o sustavima preporuke razlog je taj što s češćom interakcijom korisnika i web mjesta sustav može učiti o korisniku i poboljšavati njegov profil. Time će korisnik dobivati relevantnije preporuke i u konačnici biti zadovoljniji svojim iskustvom.

¹⁰⁷ Isto, str. 13-15.

¹⁰⁸ Ricci, F.; Rokach, L.; Shapira, B. Nav. dj., str. 5.

5. Bolje razumjeti želje korisnika – Pružatelj usluga koji koristi sustav preporuke može iskoristiti podatke o korisničkim preferencijama kako bi primijenio to znanje za poboljšavanje poslovanja na različite načine.

S korisničkog gledišta, sustavi preporuke pomažu u provedbi različitih zadataka. Neki od njih su glavni, poput preporučivanja potencijalno zanimljivih predmeta korisniku, dok su ostali sporedni.¹⁰⁹

1. Pronaći neke dobre predmete – Preporučiti korisniku neke predmete rangirane na listi uz predviđanje koliko će se svidjeti korisniku (npr. na ljestvici od jedan do pet). Ovo je glavni zadatak kojeg mnogi komercijalni sustavi izvršavaju.

2. Pronaći sve dobre predmete – Preporučiti korisniku sve predmete koji mogu zadovoljiti njegove potrebe. U takvim slučajevima nije dovoljno pronaći samo neke predmete, a to je posebice važno kada je broj predmeta relativno mali ili kada je sustav preporuke od velike važnosti, recimo u medicini ili financijama.

3. Preporučiti paket – Sustav može korisniku preporučiti skupinu predmeta koji se dobro slažu zajedno. Primjerice, korisniku koji putuje i priprema plan putovanja mogu biti preporučene različite atrakcije, destinacije i usluge smještaja na određenom području.

4. Pregledavanje – Korisnik može samo pregledavati sadržaj, bez namjere da nešto kupi. U ovom je slučaju zadatak sustava preporuke pomoći korisniku pregledavati stavke za koje je vjerojatnije da će biti zainteresiran.

5. Pronaći vjerodostojan sustav preporuke – Neki korisnici ne vjeruju sustavima preporuke i zato ih testiraju da vide koliko su dobri u davanju preporuka. Stoga neki sustavi nude određene opcije kako bi korisnici mogli testirati rad sustava i uvjeriti se u vjerodostojnost preporuka.

6. Poboljšati profil – Korisnik mora sustavu pružiti podatke o svojem ukusu i preferencijama kako bi sustav mogao stvoriti personalizirane preporuke. Ako sustav ne posjeduje nikakvo specifično znanje o korisniku, tada mu može davati samo preporuke kao i bilo kojem drugom „prosječnom“ korisniku.

7. Izraziti se – Neki korisnici nisu zainteresirani za preporuke, ali žele pridonijeti sa svojim recenzijama i ocjenama i izraziti svoje mišljenje. Ako je korisnik zadovoljan načinom na koji može izraziti svoje mišljenje, ostatak će vjeran sustavu.

¹⁰⁹ Isto, str. 6.

8. Pomoći drugima – Ako korisnici vjeruju da zajednica ima koristi od njihovog doprinosa, željet će doprinijeti ocjenjivanjem sadržaja.

9. Utjecati na druge – Korisnicima glavni cilj može biti utjecati na druge kako bi kupili određeni proizvod ili kako bi konzumirali sadržaj poput filmova, knjiga, muzike i sl.

11. 5. Personalizirano pretraživanje

Personalizirano je pretraživanje (engl. *personalized information retrieval*) metoda koja koristi podatke o korisniku kako bi poboljšala rezultate pretraživanja. Kao što je ranije navedeno, korisnik pretraživanjem često neće zadovoljiti svoju informacijsku potrebu. Upiti korisnika obično su kratki i ne govore ništa o njegovim individualnim potrebama, pa tako dvije osobe koje pretražuju identične pojmove mogu tražiti sasvim drukčije rezultate.¹¹⁰

Kako bi se pretraživanje moglo personalizirati, mora postojati mogućnost modeliranja korisnikovih preferencija i interesa. To se obično ostvaruje praćenjem i agregiranjem interakcije korisnika sa sustavom. Pod interakcijom se uglavnom podrazumijevaju prethodni upiti korisnika i analiza klikova (engl. *click-through analysis*). Ove su interakcije potom iskorištene kako bi se oblikovao korisnički profil koji se može koristiti tijekom idućih pretraživanja, a obično se koristi u dvije situacije. Prva od njih je proširivanje originalnog upita korisnika dodavanjem novih termina na temelju njegovog profila, a druga se odnosi na re-rangiranje i filtriranje dobivenih rezultata u skladu s korisnikovim profilom.¹¹¹ Profili korisnika mogu se izrađivati i tako da sami korisnici izvrše registraciju i ispune upitnik u kojem eksplicitno iznesu svoje interese. U složenijim slučajevima, u obzir se uzimaju dodatne informacije, poput razine obrazovanja korisnika i upoznatost s područjima za koja je korisnik zainteresiran.¹¹²

Međutim, ovo nije lako ostvarivo. Personalizacija rezultata pretraživanja iznimno je zahtjevan postupak. Tipična tražilica obično obrađuje stotinu upita u sekundi i pruža usluge milijunima korisnika, a proizvesti personalizirane rezultate u istom vremenu nije jednostavno.¹¹³ Sljedeći je problem što mnogi korisnici osjećaju da im je profiliranjem narušena privatnost, te strahuju kako će se ovi podaci iskoristiti za zlonamjerne radnje poput spamanja, phishinga ili izlaganja njihovih podataka javnosti. Nadalje, prethodne interakcije

¹¹⁰ Carmel, D. ... et al. Personalized social search based on the user's social network. // Proceedings of the 18th ACM conference on Information and knowledge management. New York : ACM, 2009., str. 1227.

¹¹¹ Isto, str. 1227.

¹¹² Micarelli, A. ... et al. Nav. dj., str. 199.

¹¹³ Isto, str. 197.

korisnika ne pružaju uvijek dobar pokazatelj njegovih trenutnih potreba. Ovo se pokazalo posebice točno za nove korisnike jer za njih postoje ograničene osobne informacije, te u slučajevima kada se korisničke preferencije mijenjaju tijekom vremena.¹¹⁴ Osim toga, korisnici su upoznati sa sučeljima najpoznatijih tražilica, a ako personalizirano pretraživanje sa sobom donosi i nove značajke, korisnici ih možda neće razumjeti i znati upotrebljavati, ili im se neće svidjeti i željet će ostati pri onome što im je dobro poznato.¹¹⁵

Za izradu korisničkih profila mogu se koristiti i javne aktivnosti korisnika poput korisničkog ocjenjivanja, oznaka, komentara, blogova i sl. U tom slučaju nema bojazni od narušavanja privatnosti korisnika. U slučajevima kada je korisnik član neke društvene mreže, profili njegovih prijatelja mogu se iskoristiti kako bi se predvidjele njegove preferencije, pod pretpostavkom da blisko povezane osobe imaju slične interese.¹¹⁶

Noll i Meinel u svom radu iz 2007. opisuju prototip sustava namijenjenog personalizaciji pretraživanja koji se temelji na korisničkim oznakama. Koriste pristup u kojem sustav ne modificira originalan korisnički upit, već re-rangira rezultate u skladu s personaliziranim profilom korisnika. Sustav ne ovisi o pretraživaču kojeg korisnik koristi, tako da su korisnici slobodni izabrati onaj kojeg preferiraju i inače koriste. Kada god korisnik dodijeli oznaku nekom dokumentu, sustav o njemu, kao i o tom dokumentu, saznaje nove informacije koje koristi za personalizaciju. Ovaj pristup ne traži od korisnika da ulože puno truda jer ne moraju svoje preferencije izravno unositi u sustav, a ni ne narušava njihovu privatnost jer ih nitko ne nadzire i prati njihove podatke bez da su oni svjesni toga. Korisnici točno znaju kada se i koji podaci prikupljaju za personalizaciju, a od dodjeljivanja korisničkih oznaka korisnici imaju izravnu korist bez obzira na to koriste li se one za personalizaciju ili ne.¹¹⁷

Tehnika koja su autori osmislili zove se *tagmarking*, a sastoji se od dva glavna elementa. Prvi je korak prikupljanje i agregiranje podataka o korisnicima i dokumentima, a drugi personalizacija pretraživanja na temelju tih podataka. Prikupljaju se podaci o oznakama (engl. *tags*) i spremljenim dokumentima (engl. *bookmarks*). Kada korisnik označi dokument na webu, pohranjuje ga za kasniju upotrebu. To dovodi do dvije pretpostavke: 1. korisnici

¹¹⁴ Carmel, D. ... et al. Nav. dj., str. 1227.

¹¹⁵ Micarelli, A. ... et al. Nav. dj., str. 197.

¹¹⁶ Carmel, D. ... et al. Nav. dj., str. 1228.

¹¹⁷ Noll, M. G.; Meinel, C. Web search personalization via social bookmarking and tagging. // The semantic web / edited by Aberer K. et al. Berlin : Springer, 2007., str. 367-368.

označuju samo dokumente koji su vrijedni i relevantni za njih, 2. korisnici imaju poticaj za dodavanje vrijednih metapodataka spremljenim dokumentima.¹¹⁸

Tagmarking koristi ekstenziju za preglednik koja omogućava korisnicima da „tagmarkiraju“ stranice koje su pronašli pretraživanjem. Kada korisnik upiše ključne riječi u tražilicu i provede pretraživanje, sustav mu omogućava da u njegovu memoriju pohrani upit dok on vrednuje dobivene rezultate. Primjerice, ako je korisnik pretraživao pojmove „gavran“ i „Poe“, kada pronađe relevantan dokument može ga označiti klikom na *tagmark* gumb, a sustav će automatski prevesti upit u oznake (oznake bi u ovom slučaju bile „gavran“ i „Poe“) i dodijeliti ih označenom dokumentu. Ovime se korisnicima omogućava da pohrane dokument i odmah mu dodijele oznake jednim klikom, a istovremeno se prikupljaju podaci za personalizaciju profila.¹¹⁹

Personaliziranom pretraživanju upućuju se određene kritike. Američki aktivist Eli Pariser 2010. skovao je termin filter mjehurić (engl. *filter bubble*), a njime opisuje stanje intelektualne izolacije koje nastaje personalizacijom pretraživanja jer su korisnici izloženi samo informacijama koje odgovaraju njihovim korisničkim podacima. Pariser je na uvođenje ovog koncepta potaknula Googleova inovacija iz 2009. kada je počeo koristiti različite podatke o korisnicima, poput povijesti pretraživanja, preglednika koji koriste, lokacije s koje se prijavljuju na svoje račune i sl., kako bi na temelju tih informacija stvorio korisnički profil i personalizirao rezultate pretraživanja.¹²⁰

Velikim kompanijama poput Googlea, Netflix, Facebooka, YouTubea, Amazona i sl. personalizacija je glavna strategija. Što su informacije koje nude relevantnije i osobnije za korisnika, to kompanije mogu prodati više proizvoda i reklamnog prostora. Međutim, personalizacija nije samo ciljano oglašavanje i ne utječe samo na izbor pri kupovini, već i na vrstu i količinu informacija koje dopiru do korisnika.¹²¹

Dvije osobe čiji se upit sastoji od istih ključnih riječi mogu dobiti sasvim različite rezultate. Pariser daje primjer iz 2010. kada je došlo do masovnog izljeva nafte u Meksičkom zaljevu za kojeg je bila odgovorna kompanija British Petroleum (BP). Dvije su korisnice sličnih karakteristika (obje visoko obrazovane, lijevo politički orijentirane, žive na istom području) pretraživale pojam „BP“. Prva stranica rezultata za jednu od njih sadržavala je

¹¹⁸ Isto, str. 369.

¹¹⁹ Isto, str. 370.

¹²⁰ Pariser, E. *The filter bubble: What the Internet is hiding from you*. London : Penguin Press, 2011., str. 6.

¹²¹ Isto, str. 9-10.

vijesti o izljevu nafte, dok je druga na prvoj stranici pronašla vijesti o investicijama u kompaniju. I broj rezultata bio je različit, u jednom je pretraživanju pronađeno oko 180 milijuna rezultata, a u drugom oko 139 milijuna.¹²²

Iako su ljudi i prije pratili medije i sadržaj koji ih zanima, a ignorirali ostalo, koncept filter mjehurića unio je neke novosti. Kao prvo, korisnici se u svom mjehuriću nalaze sami. Ako osoba na televiziji gleda program kojeg ne prati mnogo ljudi, ili čita novine koje nisu osobito popularne, svejedno će postojati i drugi koji gledaju ili čitaju isti sadržaj. Pariser smatra kako se nalazimo u dobu u kojemu su informacije koje dijelimo temelj zajedničkog iskustva, a filter mjehurić je sila koja nas razdvaja. Nadalje, filter mjehurić je nevidljiv. Korisnik ne može znati koliko su rezultati koje dobiva pretraživanjem pristrani (često uopće ni ne zna da su pristrani) jer ne zna na kojim osnovama je Google personalizirao rezultate, i da li su pretpostavke koje Google ima o njemu uopće točne. Naposljetku, korisnik ne odlučuje sam ući u mjehurić. Danas je teško pronaći web stranice koje ne personaliziraju rezultate, tako da će se vjerojatno svaki korisnik interneta susresti s rezultatima prilagođenima samo njemu. S druge strane, pri čitanju novina ili gledanju vijesti na televiziji moguće je barem približno procijeniti da li je riječ o pristranom sadržaju.¹²³

12. Istraživanja društvenog pretraživanja

Kako bi se ilustrirale mogućnosti primjene društvenog pretraživanja u svakodnevnom korištenju, analizirano je pet istraživanja provedenih u posljednjih deset godina. Istraživanja govore o interakcijama u koje se ispitanici upuštaju tijekom procesa pretraživanja, pretraživanju putem pretraživača koji se baziraju na nekoj vrsti društvenog pretraživanja, iskustvu i zadovoljstvu korisnika s takvim oblikom pretraživanja i uspješnosti u odnosu na tradicionalno pretraživanje.

Chi je 2009. proveo istraživanje u kojem je na 150 ispitanika ispitao da li se prije, tokom ili nakon pretraživanja upuštaju u interakciju, te kakva je priroda tih interakcija. Konkretnije, istraživao je razloge njihova posljednja pretraživanja, kako su proveli pretraživanje i da li su učinili nešto s dobivenim rezultatima. Dobiveni nalazi pokazuju kako se 43% ispitanika upustilo u društvenu interakciju prije pretraživanja weba. Korisnici su u tim interakcijama najčešće tražili savjet ili preporuku vezanu uz ključne riječi i web stranice. Ovo ukazuje da prije procesa pretraživanja postoji određena društvena priprema. Nadalje, oko 40%

¹²² Isto, str. 6.

¹²³ Isto, str. 13.

ispitanika sudjelovalo je u interakciji tijekom pretraživanja, što nije neobično, jer korisnici često provode istraživanje u paru ili grupi, pritom sugerirajući ključne riječi osobi koja pretražuje. Na kraju, pokazalo se da 60% ispitanika obično s drugima podijeli informacije koje su dobili pretraživanjem.¹²⁴

Morris i sur. 2010. proveli su istraživanje kojim su htjeli usporediti proces pretraživanja weba s traženjem informacija putem Facebooka, kvalitetu dobivenih informacija i zadovoljstvo korisnika. U istraživanju je sudjelovalo 12 ispitanika koji su postavili pitanje na svom Facebook profilu, a istodobno su pokušavali pronaći odgovor na njega pretraživanjem weba. Sudionici su morali imati barem 50 prijatelja na Facebooku, a pitanje su odabirali sami kako bi najbolje mogli procijeniti da li su dobili informacije koje su tražili.¹²⁵

Ispitanici su također sami određivali vrijeme pretraživanja na webu, završili bi pretraživanje tada kad bi bili zadovoljni nađenim informacijama, a nakon toga su pogledali kakve su odgovore dobili na svoj upit na Facebooku. Tri dana nakon istraživanja ispitanici su istraživačima poslali sliku svog profila kako bi mogli vidjeti ima li novih odgovora nakon što je istraživanje završilo. Rezultati su pokazali kako je prosječno vrijeme pretraživanja weba iznosilo 30.3 min, a ispitanici su u prosjeku unijeli 6.5 upita i posjetili 35.4 stranice. Kada su ispitanici provjerili svoje Facebook profile nakon završetka pretraživanja, u prosjeku su pronašli 1.4 odgovor na svoje pitanje, a maksimalan broj odgovora bio je 5. Broj odgovora bio je u korelaciji s brojem prijatelja na Facebooku, što su korisnici imali više prijatelja, veći je bio i broj odgovora. Nakon trećeg dana, ispitanici su u prosjeku dobili 4.1 odgovora, a broj dobivenih odgovora kretao se od 0 do 20 (dva ispitanika nisu dobila niti jedan odgovor).¹²⁶

Što se tiče zadovoljstva s dobivenim rezultatima, 91.7% ispitanika bilo je zadovoljnije rezultatima koje su dobili pretraživanjem weba, a kao prednosti takvog pristupa naglašavaju brzinu, autonomiju u upravljanju pretraživanjem, mogućnost da „skaču“ s jednog upita na drugi te mogućnost reformuliranja upita nakon što su saznali više o temi. Međutim, ispitanici su uočili i razne prednosti traženja informacija putem društvenih mreža. Ono je pogodno kada je riječ o nekom subjektivnom pitanju zato što vas prijatelji na društvenim mrežama dobro poznaju, znaju vaš ukus i interese, što daje dodatni kontekst. Također, ispitanici su izjavili kako imaju visoko povjerenje u odgovore koje dobiju od svojih prijatelja. Čak i ako odgovori

¹²⁴ Chi, E. H. Information seeking can be social. // *Computer* 42, 3(2009), str. 42-43.

¹²⁵ Morris, M. R.; Teevan, J.; Panovich, K. Comparison of Information Seeking Using Search Engines and Social Networks. // *Proceedings of the Fourth International Conference on Weblogs and Social Media*. California : The AAAI Press, 2010., str. 291.

¹²⁶ Isto, str. 292.

nisu imali informacijsku vrijednost, imali su društvenu. Postavljanje upita omogućilo je ljudima kontaktiranje s prijateljima, s kojima se možda nisu čuli duže vremena. Na kraju, ispitanici su putem Facebooka dobili neke informacije koje ne postoje na webu. Autori istraživanja zaključuju kako umjesto toga da jedna metoda pretraživanja bude superiornija, isplativije je koristiti ih zajedno, pogotovo u onim slučajevima kada su korisnici nesigurni rezultatima dobivenim pretraživanjem weba, mogu postaviti pitanje prijateljima na društvenim mrežama.¹²⁷

Slično su istraživanje 2009. proveli Evans i sur. U njihovom je istraživanju 8 ispitanika trebalo pronaći informacije o pitanju povezanom s američkom energetsom politikom, a ta je tema odabrana jer je u to vrijeme trajala žustra debata oko visokih cijena plina. Ispitanici su odgovore na pitanja tražili putem web pretraživača i društvenim pretraživanjem (blogovi, sustavi pitanja i odgovora, prim. Yahoo Answers, e-pošta, instant poruke, komunikacija licem u lice, telefonski poziv...). Autori su identificirali tri taktike koje su ispitanici koristili : 1. direktna pitanje – primjerice prijatelju ili kolegi za kojeg su smatrali da posjeduje specifično znanje, 2. javna pitanje – postavljanje pitanja na društvenim mrežama ili nekim drugim stranicama, 3. pretraživanje – pretraživanje baza društvenih podataka (npr. Yahoo! Answers, blogovi i sl.).¹²⁸

Svaka od ovih taktika ima potencijalne kvalitete. Ciljanje odabranih pojedinaca korisno je ako posjeduju znanje koje je korisnicima potrebno i ako su u tom trenutku dostupni. Postavljanje javnih pitanja obuhvaća veći broj korisnika i teoretski povećava vjerojatnost pronalaženja pojedinaca koji mogu pružiti korisne informacije. Na kraju, pretraživanje pruža velik izvor podataka, ali ograničeno je na sadržaj koji je prisutan u bazi podataka. Njihovi su zaključci o tome koja je strategija najbolja bili slični kao i u prethodno opisanom istraživanju - najproduktivnije pretraživanje proveli su oni korisnici koji su kombinirali sve navedene taktike.¹²⁹

Morris i sur. 2010. istraživali su pitanja koja korisnici društvenih mreža postavljaju na svojim profilima u obliku statusa. U istraživanju je sudjelovalo 624 ispitanika koji su imali profile na Facebooku i Twitteru. Ispitivano je koliko često ispitanici postavljaju pitanja, kakva pitanja postavljaju i kakve odgovore dobivaju. Također, autori su od ispitanika zatražili da se ulogiraju na svoje profile i kopiraju jedno od nedavno objavljenih pitanja. Osim toga,

¹²⁷ Isto, str. 293.

¹²⁸ Evans, B. M.; Kairam, S.; Pirolli, P. Nav. dj., str. 3378-3379.

¹²⁹ Isto, str. 3380.

istraživalo se i da li ispitanici odgovaraju na tuđa pitanja te koja je u tom slučaju njihova motivacija.¹³⁰

Od svih ispitanika, njih 249 moglo je pružiti primjer nedavno postavljenog pitanja, a ostali to nisu mogli ili zbog privatnih razloga ili nisu mogli pronaći traženi status. Primjeri pitanja kretali su se od smiješnih („Zašto su muškarci tako glupi?“) do kompleksnih („Digitalna kamera je upravo prestala raditi – moram ju zamijeniti danas za sutrašnje putovanje. Recite mi što da kupim/ne kupim. Ispod \$200“). Pitanja su bila uglavnom kratka, većina ih se sastojala od jedne rečenice, u upitnom obliku. Ostatak je bio u izjavnom obliku, ali je sadržavao fraze poput „trebam“ ili „zanima me“. Ispitanici su najčešće tražili preporuke ili mišljenja svojih prijatelja, odnosno subjektivnu informaciju. Nakon njih slijede pitanja gdje se tražilo činjenično znanje i objektivni odgovori, a zatim retorička pitanja čija je svrha zapravo potaknuti diskusiju, a ne dobiti konkretan odgovor. Najmanje je ispitanika tražilo usluge, davalo ponude i pozivalo na određen događaj. Gledano tematski, najviše se pitanja odnosilo na tehnologiju (računala, hardver i softver, programiranje, mobiteli, kamere i sl.), zatim zabavu (filmovi, televizija, umjetnost, knjige, sport i sl.), nakon toga dom i obitelj (djeca, kućni ljubimci, zdravlje, kuhanje i sl.), profesionalna pitanja (posao, edukacija, konferencije i sl.), a najmanje se pitanja odnosilo na određena mjesta, restorane, kafiće, aktualne događaje, kupovinu te etička i moralna pitanja.¹³¹

Većina ispitanika odgovore je tražila putem društvenih mreža jer imaju više povjerenja u odgovore koje dobiju od osoba koje poznaju. Mnogo je njih vjerovalo i kako pretraživanje nije pogodno za njihovu informacijsku potrebu (iako prethodno nisu pokušali provesti pretraživanje). Nešto manji broj ispitanika pokušao je prvo provesti istraživanje i nije uspio, stoga je pokušao na društvenim mrežama. Također, ispitanici su postavljali pitanja jer su znali da među njihovim prijateljima postoje eksperti u području koje ih zanima, zatim jer su se željeli povezati s prijateljima, te jer je ovaj način lakši nego čitanje mnoštva rezultata dobivenih pretraživanjem. Većina ispitanika, njih 69%, bilo je zadovoljno dobivenim odgovorima. Kada je riječ o motivaciji za odgovaranje na tuđa pitanja, najčešća motivacija bila je altruistične prirode. Ekspertiza ispitanika bila je sljedeći bitan faktor, bili su motivirani odgovoriti na pitanje kada su smatrali da su stručni u tom području. Također, motivirala ih je zanimljivost samog pitanja, dobri odnosi s osobom koja je postavila pitanje, višak slobodnog

¹³⁰ Morris, M. R.; Teevan, J.; Panovich, K. What do people ask their social networks, and why? : a survey study of status message q&a behavior. // Proceedings of the SIGCHI conference on Human factors in computing systems. New York : ACM, 2010., str. 1741.

¹³¹ Isto, str. 1742-1743.

vremena i mišljenje da imaju duhovit odgovor. Demotiviranima za odgovaranje osjećali su se kada nisu imali dovoljno znanja u traženom području ili ako je pitanje preintimne prirode za društvene mreže.¹³²

Morris je 2012. proveo istraživanje u kojemu je proučavao suradničke prakse u pretraživanju weba, učestalost takvih aktivnosti, informacijske potrebe koje motiviraju suradnju, metode i alate koji se primjenjuju te zadovoljstvo korisnika. Od 167 korisnika, njih 65.3% barem su se jednom upustili u suradničko pretraživanje, a dob se pokazala kao bitan prediktor - mlađi ispitanici češće su pretraživali u paru ili grupi. Na dnevnoj bazi suradnički je pretraživalo 11% ispitanika, a njih 38.5% barem jednom tjedno. Najveći broj ispitanika pretraživao je na ovaj način iz profesionalnih razloga, odnosno u slučajevima kada im je to olakšalo posao ili je posao to od njih zahtijevao. Najviše ispitanika pretraživalo je u parovima, a zatim u grupama od troje i četvero, dok su veće grupe bile rijetke. Najviše je ispitanika koristilo računala i prijenosna računala, a pretraživali su putem tražilica, ali i društvenih mreža te sustava pitanja i odgovora.¹³³

Kako popularni pretraživači ne podržavaju komunikaciju korisnika, ispitanici koji su pretraživali na odvojenim lokacijama najčešće su komunicirali putem e-pošte, zatim pozivima, SMS-om i instant porukama. Informacijsku je potrebu zadovoljilo 82.5% korisnika, a 77.9% ih je bilo zadovoljno s lakoćom suradnje. Iako su ovi postoci visoki, pokazuju da ima mjesta za napredak, što su prepoznali i sami ispitanici. Oni smatraju kako bi se pretraživanje poboljšalo kada bi svi sudionici mogli vidjeti aktivnosti ostalih te dijeliti nađene informacije.¹³⁴

13. Zaključak

U središtu djelovanja informacijske znanosti nalazi se pretraživanje informacija, aktivnost koja označava traženje informacija u informacijskim sustavima, koje su relevantne za informacijsku potrebu koja se nastoji zadovoljiti. Iako je informacijska znanost mlada disciplina, organiziranje i pohranjivanje informacija postoji od kad su ih ljudi počeli bilježiti na glinenim pločicama i papirusima. Moderan razvoj pretraživanja informacija potekao je iz bibliotekarstva i ručnog indeksiranja, a njegov je napredak ostvaren zahvaljujući razvoju računalne tehnologije. Ključni koncept u pretraživanju informacija je relevantnost. Ona

¹³² Isto, str. 1744-1746.

¹³³ Morris, M. R. Collaborative search revisited. // Proceedings of the 2013 conference on Computer supported cooperative work. New York : ACM, 2013., str. 1181-1184.

¹³⁴ Isto, str. 1185.

ukazuje na to da li su informacije koje je sustav pronašao povezane s korisničkim upitom. Ako pronađeni dokument sadrži sve pojmove iz upita, to ne mora nužno značiti da je relevantan za korisnika, jer je relevantnost subjektivan pojam. Stoga je potrebno razlikovati korisničku relevantnost, koju procjenjuje sam korisnik, i tematsku relevantnost, koja je ostvarena ako se dokument tematski podudara s upitom. Pronađeni dokumenti evaluiraju se mjerama preciznosti i odziva, kojima se procjenjuje koliko je pronađenih dokumenata relevantno, odnosno koliko je relevantnih dokumenata pronađeno. Najjednostavniji model pretraživanja informacija sastoji se od upita koji izražava informacijsku potrebu, zbirke dokumenata i pretraživača koji uparuje relevantne dokumente s upitom. Ovaj se model fokusira na sustav, a ne samog korisnika, dok kasnije razvijeni modeli u obzir uzimaju znanje koje korisnik posjeduje, njegove želje i ciljeve.

Veliki zaokret dogodio se 1990-ih nastankom i širenjem World Wide Weba, kada su sustavi za pretraživanje postali komercijalni i široko dostupni. Pretraživanjem se nisu više bavili samo knjižničari i informacijski stručnjaci, već i obični korisnici. Opseg dokumenata na webu rastao je strahovitom brzinom, što je otežalo pronalazak relevantnih dokumenata. Razvoj algoritama za rangiranje, konkretno Googleovog PageRanka olakšao je ovaj zadatak. Zahvaljujući PageRanku, prvi rezultati koje korisnici vide relevantni su za njihov upit jer ih sačinjavaju stranice na koje se referira veći broj stranica, za koje se smatra da su vjerodostojne i imaju autoritet. Rangiranje stranica na ovaj način uz sebe veže i određene nedostatke, npr. stranice koje nisu popularne teško će biti otkrivene, a možda su relevantne za korisnika. Korisnici se razlikuju svojim iskustvom u pretraživanju, znanjem o temi koju pretražuju i tehničkom stručnošću. Navedeni faktori utječu na njihovo iskustvo pretraživanja i rezultate. Neiskusni korisnici potrošit će više vremena na pretraživanje, postaviti više upita i osjećati se izgubljeno među mnoštvom rezultata, dok će napredniji korisnici imati više samopouzdanja i koristiti naprednije tehnike u postavljanju upita.

Cilj je ovoga rada bio ispitati, kontekstualizirati i kritički analizirati novu paradigmu pretraživanja informacija, tzv. društveno pretraživanje. Društveno je pretraživanje kišobran termin koji obuhvaća sve pretraživače koji koriste korisnički generiran sadržaj kako bi poboljšali rezultate pretrage. Razvojem društvenih medija i weba 2.0 korisnici su postali kreatori sadržaja, što je omogućilo društvenom pretraživanju da iskoristi društveno stvoren sadržaj. Primjene društvenog pretraživanja raznolike su, a neke od njih su korisničko označivanje, sustavi pitanja i odgovora, zajedničko pretraživanje, suradničko filtriranje i sustavi preporuke te personalizirano pretraživanje. U usporedbi s tradicionalnim modelom

pretraživanja, društveno pretraživanje nadomješta određene nedostatke. Ono u obzir uzima veze između korisnika, preferencije, interese i potrebe korisnika, kao i kontekst pretraživanja, dok su rezultati običnog pretraživanja uglavnom univerzalni. Dok tradicionalno pretraživanje pretpostavlja da su upiti korisnika i informacije na webu usklađeni i izraženi zajedničkim jezikom, društveno pretraživanje pokušava nadvladati neusklađenosti dopuštajući korisnicima da opišu dokumente na vlastiti način. Korisnici često ne znaju izraziti svoju informacijsku potrebu i oblikovati ju u upit, a imaju i praznine u znanju, kojih ponekad nisu ni svjesni. Korisnički su upiti najčešće sastavljeni od samo dvije ili tri riječi što često nije dovoljno precizno za željeni rezultat. Navedeni razlozi utječu na loše iskustvo pretraživanja i nezadovoljavanje informacijske potrebe. Društveno pretraživanje može ukloniti ovu prepreku tako što dozvoljava korisniku da svoj upit izrazi široko koliko je to potrebno, npr. postavljajući pitanje u nekom od sustava pitanja i odgovora ili na društvenoj mreži. Međutim, društveno pretraživanje također ima nedostatke i ograničeno je u nekim segmentima, te nije pogodno za sve vrste upita. Odabir vrste pretraživanja uvelike ovisi o korisnikovoj potrebi, ali čini se da najbolji učinak proizlazi iz kombinacije različitih vrsta pretraživanja.

14. Literatura

Amershi, S.; Morris, M. R. CoSearch: a system for co-located collaborative web search. // Proceedings of the SIGCHI conference on human factors in computing systems. New York : ACM, 2008. Str. 1647-1656.

Amitay, E. ... et. al. Social search and discovery using a unified approach. // Proceedings of the 20th ACM conference on hypertext and hypermedia. New York : ACM, 2009. Str. 199-208.

Baeza-Yates, R.; Ribeiro-Neto, B. Modern information retrieval. 2nd ed. Harlow : Pearson, 2011.

Bates, M. J. The design of browsing and berrypicking techniques for the online search interface. // Online review 13, 5(1989), str. 407-424.

Bouadjenek, M. R.; Hacid, H.; Bouzeghoub, M. Social networks and information retrieval, how are they converging? A survey, a taxonomy and an analysis of social information retrieval approaches and platforms. // Information Systems 56, (2016), str. 1-18.

Broder, A. A taxonomy of web search. // ACM Sigir forum 36, 2(2002), str. 3-10.

Buckley, C.; Salton, G.; Allan, J. The effect of adding relevance information in a relevance feedback environment. // SIGIR'94 / edited by Croft B.W.; van Rijsbergen C.J. London : Springer, 1994. Str. 292-300.

Burghardt, M.; Heckner, M.; Wolff, C. The many ways of searching the web together : A comparison of social search engines. // Web search engine research / edited by Dirk Lewandowski. Vol.4. United Kingdom : Emerald Group Publishing Limited, 2012. Str. 19-46.

Carmel, D. ... et al. Personalized social search based on the user's social network. // Proceedings of the 18th ACM conference on Information and knowledge management. New York : ACM, 2009. Str. 1227-1236.

Ceri, S. ... et. al. Web information retrieval. New York : Springer, 2013.

Chi, E. H. Information seeking can be social. // Computer 42, 3(2009), str. 42-46.

Croft, W. B.; Metzler, D.; Strohman, T. Search engines: Information retrieval in practice. Reading : Addison-Wesley, 2010.

Cross, R.; Rice, R. E.; Parker, A. Information seeking in social context: Structural influences and receipt of information benefits. // IEEE Transactions on Systems, Man, and Cybernetics, Part C (applications and reviews) 31, 4(2001), str. 438-448.

Evans, B. M.; Kairam, S.; Pirolli, P. Exploring the cognitive consequences of social search. //CHI'09 Extended Abstracts on Human Factors in Computing Systems. New York : ACM, 2009. Str. 3377-3382.

Goldman, E. Search engine bias and the demise of search engine utopianism. // Web Search : Multidisciplinary Perspectives / edited by Spink, A.; Zimmer, M. Berlin : Springer, 2008. Str. 121-133.

Golovchinsky, G.; Pickens, J.; Back, M. A taxonomy of collaboration in online information seeking. // Search 32, (2009), str. 1-3.

Horowitz, D.; Kamvar, S. D. The anatomy of a large-scale social search engine. // Proceedings of the 19th international conference on World wide web / edited by Rappa, M. New York : ACM, 2010. Str. 431-440.

Ingwersen, P. Information retrieval interaction. London : Taylor Graham, 1992.

McDonnell, M.; Shiri, A. Social search: A taxonomy of, and a user-centred approach to, social web search. // Program 45, 1(2011), str. 6-28.

Micarelli, A. ... et al. Personalized search on the world wide web. // The adaptive web : Methods and Strategies of Web Personalization / edited by Brusilovsky, P. et. al. Berlin : Springer, 2007. Str. 195-230.

Morris, M. R. Collaborative search revisited. // Proceedings of the 2013 conference on Computer supported cooperative work. New York : ACM, 2013. Str. 1181-1192.

Morris, M. R.; Horvitz, E. SearchTogether: an interface for collaborative web search. // Proceedings of the 20th annual ACM symposium on User interface software and technology. New York : ACM, 2007. Str. 3-12.

Morris, M. R.; Teevan, J.; Panovich, K. Comparison of Information Seeking Using Search Engines and Social Networks. // Proceedings of the Fourth International Conference on Weblogs and Social Media. California : The AAAI Press, 2010. Str. 291-294.

Morris, M. R.; Teevan, J.; Panovich, K. What do people ask their social networks, and why? : a survey study of status message q&a behavior. // Proceedings of the SIGCHI conference on Human factors in computing systems. New York : ACM, 2010. Str. 1739-1748.

Noll, M. G.; Meinel, C. Web search personalization via social bookmarking and tagging. // The semantic web / edited by Aberer K. et al. Berlin : Springer, 2007. Str. 367-380.

Pariser, E. The filter bubble: What the Internet is hiding from you. London : Penguin Press, 2011.

Pretraživanje informacija. // Hrvatska enciklopedija. Mrežno izd. Zagreb : Leksikografski zavod Miroslav Krleža, cop. 2019. Dostupno na: <http://www.enciklopedija.hr/natuknica.aspx?id=50252> (16.8.2019.).

Previšić, J.; Ozretić Došen, Đ. Internetski marketing. Obnovljeni separat poglavlja 19. u knjizi. Zagreb : Marketing, Adverta, 2004.

Ricci, F.; Rokach, L.; Shapira, B. Introduction to recommender systems handbook. // Recommender systems handbook / edited by Rokach, L. et. al. Boston : Springer, 2011. Str. 1-35.

Richman, D. Social search comes of age. // Information Outlook 11, 8(2007), str. 1-11.

Rose, D. E.; Levinson, D. Understanding user goals in web search. // Proceedings of the 13th international conference on World Wide Web. New York : ACM, 2004. Str. 13-19.

Russell-Rose, T.; Tate, T. Designing the search experience: The information architecture of discovery. Amsterdam : Morgan Kaufmann, 2013.

Sahami, M. ... et al. The happy searcher: Challenges in web information retrieval. // Pacific Rim International Conference on Artificial Intelligence / edited by Zhang, C. et. al. Berlin : Springer, 2004. Str. 3-12.

Sanderson, M.; Croft, W. B. The history of information retrieval research. // Proceedings of the IEEE 100 (Special Centennial Issue), (2012), str. 1444-1451.

Smyth, B.; Coyle, M.; Briggs, P. Communities, collaboration, and recommender systems in personalized web search. // Recommender Systems Handbook / edited by Rokach, L. et. al. Boston : Springer, 2011. Str. 579-614.

Terveen, L.; Hill, W. Beyond recommender systems: Helping people help each other. // Human-Computer Interaction in the New Millennium / edited by Carroll, J.M. 1st ed. New York : ACM Press, 2001. Str. 1-21.

Vuorikari, R.; Koper, R. Ecology of social search for learning resources. // Campus-Wide Information Systems 26, 4(2009), str. 272-286.

Društveno pretraživanje: primjena, mogućnosti i izazovi

Sažetak

Razvoj društvenih medija i weba 2.0. omogućio je iskorištavanje društveno stvorenog sadržaja kako bi se poboljšali rezultati pretraživanja. U radu se prikazuju mogućnosti primjene društvenog pretraživanja, njegove prednosti i nedostaci. Cilj je pokušati odgovoriti na pitanje da li se društveno pretraživanje može uspješno koristiti kao nadopuna tradicionalnom pretraživanju. Tradicionalno pretraživanje često ostavlja korisnike nezadovoljnima jer nisu uspjeli pronaći odgovor na svoju informacijsku potrebu. Nekada se to događa zbog samih korisnika koji ne znaju pravilno oblikovati svoju potrebu u upit, a nekada zbog toga što pretraživači zanemaruju kontekst pretraživanja i specifične potrebe pojedinog korisnika. Društveno pretraživanje pokušava razriješiti oba navedena problema.

U prvom je dijelu rada opisan pojam pretraživanja informacija i objašnjen način na kojem funkcionira tradicionalno pretraživanje. Navedene su vrste modela pretraživanja, korisničkih upita i podjela korisnika prema stručnosti pretraživanja, te su analizirani utjecaji koje je razvoj World Wide Weba imao na pretraživanje informacija. U drugom dijelu rada prikazane su različite primjene društvenog pretraživanja, a kroz primjere objašnjen je način njihovog funkcioniranja i arhitektura sustava.

Ključne riječi: društveno pretraživanje, pretraživanje informacija, informacijska potreba, korisnički upit, relevantnost

Social search: application, possibilities and challenges

Summary

Social Media and development of Web 2.0. made it possible to make use of socially-generated content to improve search results. The paper presents the possibilities of social search, its advantages and disadvantages. The goal is to try to answer whether social search can be successfully used to complement traditional search. Traditional search often leaves users dissatisfied because they have been unable to find an answer to their information need. Sometimes this happens because the users do not know how to properly convey their need into query, and sometimes because search engines ignore the search context and specific needs of the individual user. Social search is trying to solve both of these problems.

The first part of the paper describes the concept of information retrieval and explains how traditional search works. Types of search models, user queries, and user taxonomy are listed, and the impacts that the development of the World Wide Web has had on information retrieval have been analyzed. In the second part of the paper, various applications of social search are presented, and through the examples system architecture and the way they function is explained.

Key words: social search, information retrieval, information need, user query, relevancy