

UNIVERZITET U BANJA LUCI
PRIRODNO-MATEMATIČKI FAKULTET

KAKO RADI GOOGLE SEARCH?

(Seminarski rad iz kursa Informacione tehnologije i društvo)

Mentor:

prof. dr Dragan Matić

Student:

Dijana Krajišnjik

21/19

Banja Luka, april 2020.

SADRŽAJ:

1. UVOD.....	1
2. GOOGLE SEARCH.....	2
2.1. Istorija Google pretrage.....	2
3. ARHITEKTURA VEB PRETRAŽIVAČA	4
3.1. Pretraživač veb stranica (Web crawler – spider) – Googlebot.....	5
3.2. Indeksiranje (Indexing)	7
3.3. Rangiranje stranica i algoritmi za pretragu	8
3.4. Upiti u pretrazi – Web search query.....	10
4. ZAKLJUČAK.....	11
5. LITERATURA	12

1. UVOD

Tema ovog rada je pitanje „Kako radi Google search?“, to jest, kako Google uspijeva za manje od sekunde da pronađe nekad i preko milion odgovora i da ih sortira, tako da nam se prvo prikažu najrelevantniji u tom trenutku, bez obzira kakav upit da mu zadamo. Google search je danas jedan od najpoznatijih i najkorišćenijih internet pretraživača. Iza tog uspjeha nalazi se mnogo kompleksnih algoritama i baza za čuvanje podataka. Prvi koraci, u tom kompleksnom sistemu, su pronalaženje i označavanje svih veb stranica na internetu, koji su u Google pretrazi implementirani pomoću Googlebot-a i sistema za indeksiranje stranica. Njihova je uloga da pronađu, sačuvaju i na određeni način „razumiju“ sadržaj stranice. Sledeći značajan proces sa kojim ćemo se detaljnije upoznati je sortiranje ili rangiranje dobijenih rezultata. Ono je implementirano pomoću više kompleksnih algoritama koji razmatrajući mnoge faktore uspijevaju da odrede vrijednost svake stranice i njenu povezanost sa upitom i na osnovu toga odrede koje stranice bi najbolje odgovarale na naš upit. Pokušavajući da pronađe odgovor na upit Google nailazi na mnogo problema, kao što su, na primjer, razumijevanje svih jezika ili nepotpuni ili višeznačni upiti. U rješavanju svih ovih, i još mnogih drugih problema Google koristi i vještačku inteligenciju i mašinsko učenje, tako da se njegov rad iz dana u dan unapređuje učeći o nama i svijetu u kome živimo.

2. GOOGLE SEARCH

Gugl pretraživač (eng. Google search) je veb pretraživač koji je razvila kompanija Google 1997. godine. On je najkorišćeniji pretraživač veba na svetskoj komunikacionoj mreži na svim platformama vršeći više od 3,5 milijardi pretraga svaki dan.

Google search kao pretraživač veba (engl. web search engine) predstavlja internet servis, čija je svrha traženje informacija na vebu, i to najčešće zadavanjem ključnih riječi, a mnogo rjeđe odabirom ponuđenih stavki. Ishod pretrage se najčešće prikazuje kao spisak veb stranica koje sadrže traženu informaciju, uz mogućnost da se veb-stranice koje su odgovor na upit posjete sa strana pretraživača.

Redosled rezultata pretrage koji Google vraća se zasniva na više kompleksnih algoritama koji će kasnije biti detaljnije opisani. Google pretraga pruža i različite opcije za prilagođenu pretragu koristeći simbole za uključivanje, isključivanje, navođenje ili zahtevanje određenog ponašanja pri pretrazi i nudi specijalizovana interaktivna iskustva, kao što su status leta i praćenje paketa, vremenske prognoze, konverzija valuta, jedinica i vremena, definicije reči i drugo.

2.1. Istorija Google pretrage

Google-ova istorija počinje 1996. kada su Larry Page i Sergey Brin, studenti Univerziteta u Stanfordu, počeli sa radom na pretraživaču zvanom BackRub koji je kasnije 1997. preimenovan u Google. Riječ Google je nastalo od riječi googol koja označava broj koji se sastoji od jedinice poslije koje slijedi 100 nula. Larry je 1998. pokrenuo časopis „Google Friends newsletter“ da informiše zainteresovane o radu svoje kompanije, koji je kasnije zamijenjen sa blogom kao što je Google+. U avgustu 1998. je kompanije Google dobila svog prvog investitora Andy Bechtolsheim, koji je uložio 100 000 dolara. Tada se počinje razvijati u pravu privatnu kompaniju.

Već u maju 2000. Google pretraživač je pokrenut, pored engleskog, na još deset jezika: francuskom, njemačkom, italijanskom, švedskom, finskom, španskom, portugalskom, holandskom, norveškom i danskom. Za veoma kratko vrijeme, iste godine postaje pretraživač sa najvećom bazom indeksa na svijetu.

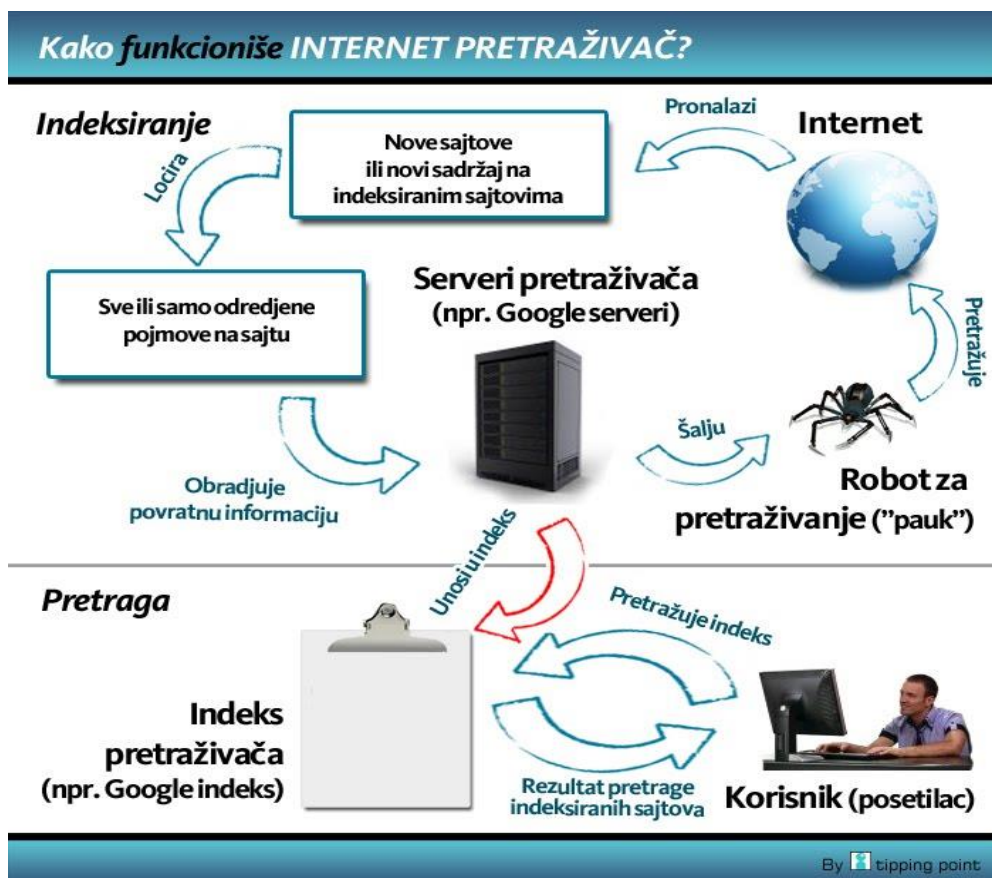
Google pretraživač nastavlja da raste, povećava broj dostupnih jezika. Uvode se razne aplikacije, kao što su Google Toolbar i Google-ova pretraga slika. Za manje od dvije godine, već krajem 2001, Google-ova baza veb indeksa se povećala čak 3 puta, dostigavši cifru od 3 milijarde veb dokumenata. U februaru 2004. taj broj se povećava do 6 milijardi, uključujući 4,28 milijardi veb stranica i 880 miliona slika, da bi u novembru iste godine dostigao osam milijardi.

U međuvremenu pokreću se razne usluge, poput Google vijesti, koja pruža korisnicima najaktuelnije članke, Frugl, danas Product Search, uslugu objektivne pretrage proizvoda koji se prodaju putem Interneta, kao i veoma popularna Gmail aplikacija, tj. besplatni email servis i Google prevodilac (internet prevodilac), kao i pretragu putem mobilnih telefona. Do kraja 2005. godine slijedi niz novih usluga koje pruža Google-ova kompanija, a to su Google Scholar, Google desktop pretragu, pretragu knjiga, videa i unapređenu pretragu slika i preteču današnjeg Google Go.

Do godine 2008. Google pretraživač je ubedljivo najtraženiji pretraživač, a veliki broj korisnika privlače i razna takmičenja, glasanja, ali i novi proizvodi, najznačajniji je pregledač Google Chrome. Popularnosti su doprineli i prerage sajtova, blogova i društvenih mreža, ali i personalizovane opcije pretrage. Google prodavnica je nastala 2010. godine, i pokazala se kao veoma uspešna, a u junu 2011. dolazi do značajnog unapređenja uvođenjem glasovne pretrage i pretrage prema slici.

3. ARHITEKTURA VEB PRETRAŽIVAČA

Arhitektura pretraživača veba se sastoji od slijedeće četiri komponente: pretraživača lokacije veb stranica (eng. web crawler), indeksera (eng. indexer), indeksa baze podataka i mašine za upite (eng. query engine). Pretraživač lokacije veba, takođe poznat kao program tragač (web spider) ili veb-robot, prolazi kroz veb u potrazi za veb stranicama slijedeći njihove URL adrese. Indeksirer je zadužen za analizu teksta svake pronađene veb stranice i vađenje ključnih riječi na osnovu kojih se onda pravi indeksna baza podataka svih analiziranih veb stranica. Kada korisnik pošalje upit, mašina za upite pretražuje indeksnu bazu podataka u potrazi za stranicama koje odgovaraju ključnim riječima koje su predmet korisnikovog upita, zatim se ti rezultati sortiraju pomoću algoritama za sortiranje i prikazuju korisniku kao niz linkova ka veb stranicama koje sadrže odgovor na poslati upit. Ovaj proces je ukratko opisan na slici koja slijedi. O svim ovim komponentama i procesima detaljnije u nastavku.



Slika 1. Prikaz načina funkcionisanja pretraživača

3.1.Pretraživač veb stranica (Web crawler – spider) – Googlebot

Prva faza u radu veb pretraživača je otkrivanje i pronalaženje svih stranica koje postoje na vebu. Ne postoji centralna baza podataka koja sadrži informacije o svim veb stranicama, pa Google mora konstantno da traži nove stranice i dodaje ih na listu poznatih stranica. Ovaj proces se naziva pretraživanje linkova veb stranica (eng. Crawling).

Googlebot je ime za Google-ov pretraživač veb stranica (eng. Crawler ili Spider). Pretraživač veb stranica je računarski program koji preuzima veb stranice sa veb servera. URL adresa svake veb stranice identifikuje njenu lokaciju na vebu. Obzirom na postojanje URL adrese, svaka veb stranica se može preuzeti sa veb-servera upotrebom protokola za prenos hiperteksta (HTTP). Polazeći od jedne početne URL adrese, pretraživač lokacija veba neprestano preuzima veb stranice bazirajući se na njihovim URL adresama i vadi URL adrese iz onih već preuzetih, kako bi pronašao i preuzeo nove. Ovaj proces se završava kada je neki uslov za zaustavljanje programa ispunjen. Neki od mogućih uslova za njegovo zaustavljanje su da nema više novih URL adresa na lokaciji ili da je unaprijed definisani broj, odnosno spisak veb stranica, već preuzet. Pošto pretraživač veb lokacija može biti u interakciji sa raznolikim samostalnim veb serverima, bitno je da bude projektovan tako da ga je lako prilagoditi novim zahtjevima.

Pretraživač veba počinje svoj rad tako što pretražuje stranice sa liste poznatih URL adresa, identifikuje sve linkove sa tih stranica i dodaje ih na listu linkova koje treba pretražiti. Takva pretraga se odvija rekurzivno poštujući nekoliko setova pravila o čemu će više biti riječ u sledećem odlomku. Prolazeći tako kroz veb stranice on sprema informacije o njima i smiješta ih u arhivu veb stranica. Arhiva predstavlja skladište koje je dizajnirano tako da može da čuva i upravlja kolekcijama veb stranica. Tu se čuvaju samo HTML dokumenti veb stranice i to najnovije verzije pronađene pomoću pretraživača. Pretraživač takođe posjeduje poseban sistem kojim sprečava beskonačno čuvanje istih veb stranica.

Ponašanje veb pretraživača je uslovljeno sledećim setovima pravila:

- pravila selekcije,
- pravila o ponovnom posjećivanju,
- pravila učtivosti,
- pravila paralelizacije,

Pravila selekcije određuju mjeru važnosti među veb stranica. Parametri pomoću kojih se to određuje su kvalitet same stranice, njena popularnost gledana kroz broj linkova ili posjeta, ili čak sama URL adresa stranice (gleda se domen stranice).

Pravila o ponovnom posjećivanju su zadužena da prate ažuriranja veb stranica i samim tim i ažuriraju podatke o njima. Dva glavna parametra koja određuju rad ovih pravila su svježina (eng. freshness) i starost (eng. age). Svježina određuje da li je stranica mijenjana ili ne u toku nekog vremena, dok starost određuje vrijeme koje je prošlo od kad je stranice zadnji put mijenjana. Cilj pretraživača je da prosječna svježina stranica bude što veća, a da prosječna starost bude što manja vrijednost. Dva osnovna pravila prema kojima se može vršiti ponovno posjećivanje su:

- pravilo uniformnosti – ono uključuje ponovno posjećivanje svih veb stranica u određenom vremenskom intervalu,
- pravilo proporcionalnosti – ono uključuje češće ponovno posjećivanje stranica koje se češće mijenjaju, tj. posjećenost je direktno proporcionalna izmjenama na veb stranici.

Pravila učtivosti služe da bi riješila problem mogućeg zagušivanja servera u slučaju da istovremeno i veb pretraživač skida veb stranice i posjetioci žele da ih koriste. Rješenja koja se koriste su uvođenje intervala koji će da ograniči veb pretraživač i tako spriječi gušenje servera njegovim zahtjevima i posebni protokol koji određuje koji dio veb stranice nije posjećen od strane pretraživača.

Pravila paralelizacije služe da bi se izbjeglo skidanje iste veb stranice više puta od strane različitih pretraživača. Ova pravila koordinišu rad više različitih pretraživača kada pretražuju isti dio veba. Glavni cilj je maksimalno uvećati brzinu preuzimanja stranica i spriječiti preuzimanje istih stranica više puta.

Kako bi se ubrzao proces pretraživanja, upotrijebljava se više preraživača lokacija veba. Oni mogu biti centralizovani i distribuirani. Centralizovani pretraživači lokacija veba se nalaze na istoj lokaciji sa koje se pokreću paralelno i sa više mašina. Distribuirani pretraživači lokacija veba se nalaze na više različitih lokacija na internetu i kontrolišu se putem jednog centralnog koordinatora, tj. svaki pretraživač lokacija veba preuzima samo veb stranice koje su mu geografski blizu. Najznačajnija prednost distribuiranih pretraživača lokacija veba tiče se smanjenja troškova komunikacija koji su rezultat njihove aktivnosti. Međutim centralizovani pretraživači su lakši za implementaciju i kontrolu nego distribuirani.

Rast i konstantne promjene na webu, stvaraju potrebu da pretraživači lokacija veba regularno vrše nova pretraživanja kao i da održavaju indeksnu bazu podataka ažuriranom. Međutim, suviše često pretraživanje veba bi dovelo do značajnog trošenja resursa, ali i poteškoće u radu veb-servera na kojima se tražene veb stranice nalaze. Stoga je potrebno upotrijebiti jednu strategiju inkrementalnog pretraživanja. Jedna od njih je da se samo pretražuju veb stranice čiji se sadržaj ili URL adresa promijenila od posljednje pretrage. Druga strategija je da se upotrijebe pretraživači lokacija veba koji imaju predefinisanu oblast pretrage, ili predefinisani skup tema za pretragu. Ovi posljednji se mogu iskoristiti za kreiranje specijalizovanih pretraživača lokacija veba koji su jedino zainteresovani za veb stranice određene tematike.

3.2.Indeksiranje (Indexing)

Google nakon što otkrije stranicu pokušava da shvati o čemu je ona. Taj proces se naziva indeksiranje (eng. Indexing). Google analizira sadržaj strane, slike i video fajlove i sve ostalo što bi mu pomoglo da razumije stranicu. Sve te informacije smješta u „Google index“ tj. ogromnu bazu podataka smještenu na mnogo servera.

Veb indeksiranje predstavlja metod indeksiranja veb stranica ili interneta u cjelini. Indeksiranje organizuje informacije o veb stranicama i smješta ih u baze na način da bi im se kasnije vrlo brzo moglo pristupiti pomoću upita i na taj način dobiti najrelevantniji mogući odgovori u što kraćem vremenskom roku. Veb pretraživač za indeksiranje koristi, na primjer, ključne riječi ili meta podatke koji i imaju ulogu da što tačnije opišu sadržaj veb stranice. Pri indeksiranju se procesuiru i sam tekst stranice kao i sav ostali sadržaj (npr. slike se mogu indeksirati pomoću teksta iz alt taga).

Pretraga svih veb stranica pojedinačno prema ključnim riječima ili naslovima bi bila jako spora, pa zato pretraživači koriste inverzne indekse. Inverzni indeksi su sistem koji koristi bazu podataka tekstualnih elemenata (pojmovi ili riječi) koji su upareni sa pokazivačima na veb stranice koje sadrže te elemente. Takođe, pretraživač koristi proces zvan tokenizacija koji redukuje riječi na njihovo osnovno značenje, tako smanjujući količinu resursa potrebnih da bi se podaci skladištili ili pretraživali. Taj sistem je mnogo brži i efikasniji od pretraživanja svakog dokumenta pojedinačno i traženja odgovarajućih riječi koje se pretražuju unutar njih. Prije dodavanja URL adresa veb stranica u bazu inverznih indeksa stvara se lista riječi za svaku veb stranicu, tj. lista koja sadrži sve svedene riječi te stranice, koja se onda transformiše i dodaje u inverznu listu. Na ovakav način se ubrzava i olakšava unošenje i ažuriranje podataka u inverznu listu.

Vrlo važan proces u indeksiranju je i već pomenuta tokenizacija. Tokenizacija je proces prevođenja sekvence karaktera ili neke druge segmente dokumenta (npr. video zapisa) u tokene tj. unaprijed poznate riječi (segmente) koji imaju definisano značenje. U toku tokenizacije mogu se javiti različiti leksički problemi kao što su: raspoznavanje jezika koji se

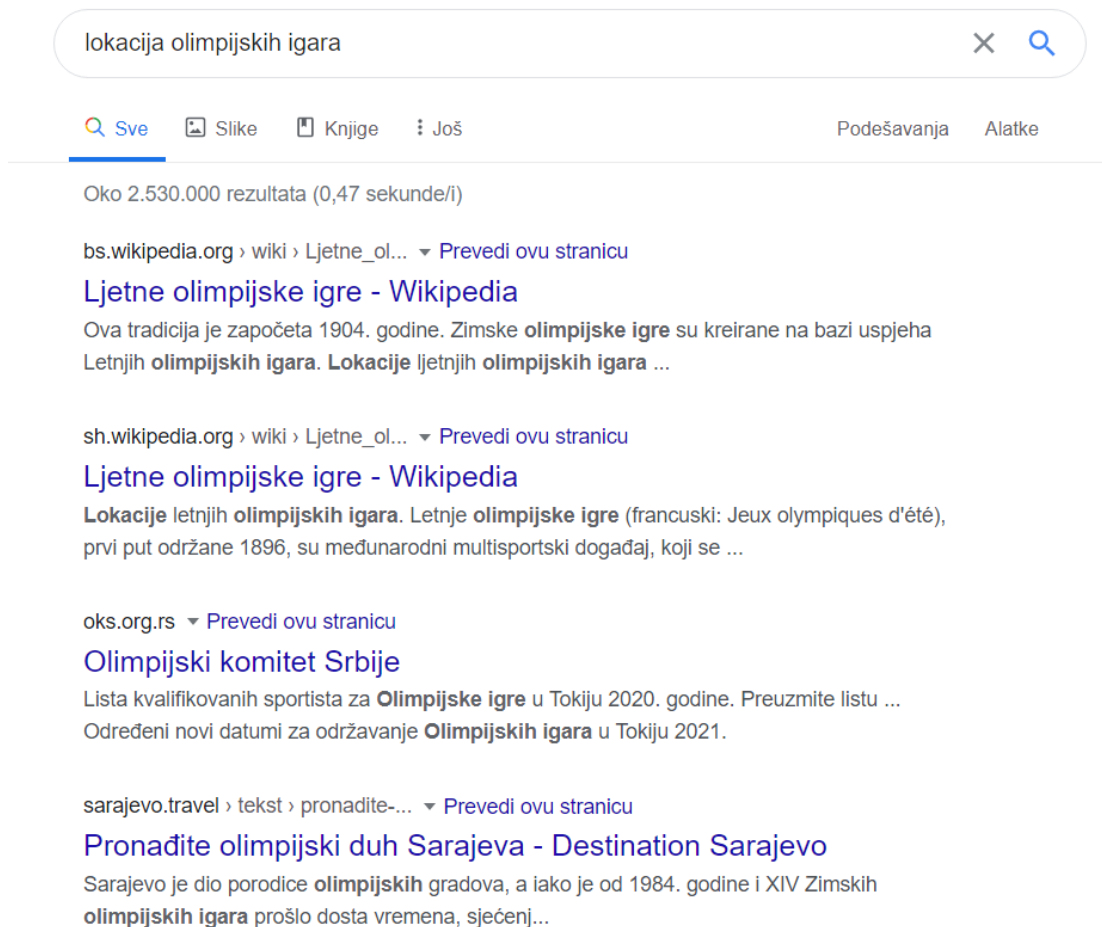
koristi u pretrazi, teško razumijevanje nekih riječi iz jezika koji su jako komplikovani za pisanje, tj. koriste posebna pisma kao što je, na primjer, kineski ili indijski, mnogo različitih gramatičkih kategorija koje postoje u različitim jezicima i naglascima (npr. padeži ili lica) itd. Pored leksičkih postoji i problem prepoznavanja različitih formata dokumenata. Tako se na primjer kod HTML dokumenata mora praviti razlika između tagova i sadržaja. Zbog svih ovih problema, ali i mnogih drugih sitnijih poteškoća, pretraga ne može postati savršena, ali već do sada je napredovala do nivoa da zadovoljava zahtjeve korisnika i svakodnevno nastavlja da se usavršava.

3.3.Rangiranje stranica i algoritmi za pretragu

Kada korisnik unese neki upit za pretragu, Google pokušava da pronađe najrelevantniji odgovor prema njegovoj bazi indeksa i to na osnovu mnogih faktora (preko 200). Neki od osnovnih faktora koji određuju kakav odgovor će biti su: poklapanje riječi iz upita sa stranicom, naša lokacija, jezik, istorija pretrage i uređaj koji koristimo (računar ili telefon). Na primjer, pretraživanjem „prodavnica tehnike i računara“ dobiće se različiti rezultati ako se pretražuje iz različitih država. Da ne bismo dobijali rezultate koji nas ne zanimaju na prvim stranicama zaslužno je rangiranje stranica. Za rangiranje stranica Google koristi mnogo algoritama od kojih su najpoznatiji PageRank, RankBrain i QualityRating.

PageRank je algoritam u Google pretrazi koji služi da prije prikazivanja rangira rezultate pretrage. PageRank je ustvari način mjerenja i određivanja važnosti veb stranice. PageRank radi na principima brojanja linkova koji pokazuju na veb stranicu i određivanju njihove važnosti. Na primjer, kada na neku stranicu vodi mnogo linkova sa različitih izvora ona se uzima kao značajna, a linkovi prema drugim stranicama koji se nalaze na takvoj stranici imaju veću vrijednost i time se povećava kvalitet i tih stranica na koje ona pokazuje.

RankBrain je algoritam za sortiranje rezultata pretrage koji koristi mašinsko učenje i vještačku inteligenciju da odredi najrelevantnije rezultate. On se sastoji od dvije faze. Prva faza određuje sve rezultate koji su odgovor na upit. Druga faza je da upit prolazi kroz „model za razumijevanje“ koji uzima u obzir faktore kao što su lokacija onoga ko pretražuje, njegova prethodna interesovanja i neke druge informacije koje su do sada prikupljene o njemu na osnovu prethodnih pretraga, riječi koje se pojavljuju u upitu i koje određuju njegovu namjeru. Cilj ovog algoritma je da što više „nauči“ o onome ko pretražuje da bi mogao da shvati i predvidi koji rezultati bi najbolje zadovoljili korisnika. Na primjer, kada u polje za pretragu otkucamo „lokacije olimpijskih igara“ pretraživač treba da otkrije šta mi stvarno želimo da saznamo, pa se tu postavljaju razna pitanja kao npr. Da li se misli na zimske ili ljetne olimpijske igre? Da li se odnosi na olimpijske igre koje su u toku ili koje su ranije završene ili koje tek treba da budu? Da li pokušavamo da dođemo do lokacije gdje se trenutno održavaju olimpijske igre? Želimo li da saznamo lokaciju prvih olimpijskih igara u Grčkoj? Da bi se odgovorilo na sva ova pitanja pretraživač mora da uzme u obzir sve dostupne informacije koje bi mu pomogle, pa npr. ako se osoba koja pretražuje nalazi u gradu gdje se trenutno održavaju olimpijske igre, ona bi vjerovatno željela da zna kako doći do te lokacije, pa bi odgovor mogao da bude ruta na Google mapi koja vodi do te lokacije. Kao što se vidi na primjeru slike ispod za nas bi to bile osnovne informacije o ljetnim olimpijskim igrama sa vikipedije i neke stranice vezane za olimpijske igre u Sarajevu.



Slika 2. Izgled pretrage „lokacija olimpijskih igara“

QualityRating je sistem rangiranja koji za osnovu vrednovanja stranice uzima njen kvalitet, tj. koliko dobro stranica ispunjava svoju svrhu. Zbog postojanja različitih tipova veb stranica, postoji i različiti kriteriji za određivanje njihovog kvaliteta. Neke od karakteristika visoko kvalitetnih stranica su:

- visok nivo ekspertinze, autoritativnosti i pouzdanosi – Expertise, Authoritativeness, and Trustworthiness (E-A-T),
- zadovoljavajuća kvaliteta naslovne stranice, uključujući i opise i korisne naslove,
- zadovoljavajuće informacije o veb stranici ili o onome ko je odgovoran za nju (npr. stranice koje služi za prodaju ili uključuju novčane transakcije treba da imaju zadovoljavajući korisnički servis),
- pozitivna reputacija kreatora stranice ili njenog vlasnika.

3.4. Upiti u pretrazi – Web search query

Upiti u pretrazi su upiti bazirani na specifičnim terminima u pretrazi koje je korisnik unio u veb pretraživač da bi dobio odgovarajuću informaciju. To je obično tekst ili hipertekst sa opcionalnim direktivama za pretragu. Postoje tri kategorije upita koje pokrivaju većinu veb pretraga, a to su:

- informativni upiti – upiti koji pokrivaju osnovne teme (npr. računar ili Banja Luka) za koje mogu da postoje hiljade odgovarajućih rezultata,
- navigacijski upiti – upiti koji pokazuju na pojedinačnu veb stranicu (npr. youtube ili pmfunibl),
- transakcijski upiti – upiti kojima korisnik pokušava da izvede neku radnju, poput kupovine automobila ili preuzimanja neke aplikacije.

Google dozvoljava upotrebu specijalnih simbola ili riječi da bi pobliže definisali našu pretragu. Ti specijalni simboli i riječi se nazivaju operatori pretrage. Neke od češćih operatora i tehnika koje se koriste u pretrazi su sledeće:

- pretraga socijalnih medija – postavlja se znak @ ispred riječi da bi se pretražili socijalni mediji, npr. @twitter,
- pretraga prema cijeni – postavlja se znak \$ ispred broja, npr. kamera \$400,
- pretraga heštegova – postavlja se znak # ispred riječi, npr. #ostanikodkuce,
- isključivanje riječi iz pretrage – postavlja se znak - ispred riječi koja treba biti izostavljena, npr. jaguar brzina -automobil,
- pretraga prema potpunom preklapanju – riječi ili fraze se pišu unutar navodnika, npr. „najviša građevina“,
- pretraga u određenom rangi brojeva – postavlja se .. između dva broja, npr. kamera \$50..\$100,
- kombinovana pretraga – postavljanje riječi „OR“ između upita, npr. maraton OR utrka,
- pretraga određene veb stranice – postavljanje „site:“ ispred imena stranice ili domena, npr. site:youtube.com,
- pretraga za povezanim (sličnim) veb stranicama – postavlja se „related:“ ispred veb adrese stranice koja je već poznata, npr. related:time.com
- pregled Google-ove keširane verzije sajta – postavljanje „cache:“ ispred adrese veb stranice

4. ZAKLJUČAK

Google search je najzastupljeniji internet pretraživač današnjice jer je uspio da organizuje i manipuliše sa milijardama veb stranice širom svijeta koje se svakodnevno mijenjaju. Njegov rad je zasnovan na osnovnim mehanizmima za pretraživanje kao što su pretraživanje linkova veb stranica, koje ostvaruje pomoću Googlebot-a, indeksiranje stranica, razumijevanje upita i rangiranja rezultata pretrage. Google search pomoću ovih mehanizama, koji su zasnovani na ogromnim sistemima algoritama, uspijeva u djeliću sekunde da odgovori na svaki upit. Bez njega današnji svijet kakvog poznajemo ne bi bio isti, tj. bez obzira na postojanje svih informacija ne bi nam bile dostupne i bilo bi jako teško doći do njih.

5. LITERATURA

1. <https://sr.wikipedia.org/sr-el/Веб-претраживач>
2. https://sr.wikipedia.org/sr-el/Гугл_претрага
3. <https://support.google.com/webmasters/answer/70897?hl=en>
4. <https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchquality/evaluatorguidelines.pdf>
5. <https://www.google.com/intl/hr/search/howsearchworks/>
6. https://en.wikipedia.org/wiki/Web_crawler
7. <https://www.deepcrawl.com/knowledge/technical-seo-library/search-engine-indexing/>
8. https://en.wikipedia.org/wiki/Search_engine_indexing
9. <https://en.wikipedia.org/wiki/PageRank>
10. <https://moz.com/learn/seo/google-rankbrain>
11. https://en.wikipedia.org/wiki/Web_search_query