

MOGUĆNOSTI PRIMENE BIG DATA KONCEPTA U DRUŠTVENIM MREŽAMA**POSSIBILITIES OF APPLYING THE BIG DATA CONCEPT IN SOCIAL NETWORKS**Ana Petković, *Fakultet tehničkih nauka, Novi Sad***Oblast – TELEKOMUNIKACIONI SISTEMI**

Kratak sadržaj – *Ovaj rad se bavi analizom Big Data tehnologije, uvidom u mogućnosti Hadoop platforme za menadžment podataka u distribuiranom okruženju, opisom prirode i značaja društvenih mreža za razvoj savremenog društva, mogućnostima primene ove tehnologije u različitim sferama društvenog života, dok se poseban naglasak stavlja na mogućnosti i benefite primene Big Data u društvenim mrežama, kao i analizom podataka istraživanja o mišljenjima ispitanika o Big Data tehnologiji.*

Ključne reči: *Big Data, Hadoop, društvene mreže, primena Big Data tehnologije*

Abstract – *This paper is about the analysis of Big Data technology, insight into the possibilities of Hadoop platform for data management in a distributed environment, description of the nature and importance of social networks for the development of modern society, possibilities of application of this technology in various spheres of social life, and about possibilities and advantages of applying Big Data on social networks and analysis of research data on the opinions of respondents about Big Data technology.*

Keywords: *Big Data, Hadoop, social networks, application of Big Data technology*

1. UVOD

Cilj samog rada jeste da pokaže šta je to Big Data tehnologija, njene karakteristike, kao i to koliko je značajna njena generalna primena u savremenom društvu, posebno primena u društvenim mrežama.

U radu je opisan koncept Big Data tehnologije, njen nastanak, njeni sastavni elementi, kao što su Hadoop i MapReduce. Ono što je takođe bilo bitno istaći su društvene mreže i njihov uticaj na današnje doba i povezanost sa Big Data tehnologijom. Naravno, tu je i generalna primena ove tehnologije, a za kraj, dat je prikaz u obliku ankete o tome šta ispitanici misli i znaju o Big Data tehnologiji.

2. BIG DATA TEHNOLOGIJA

Digitalna revolucija, odnosno razvoj Interneta, inteligentnih uređaja, društvenih mreža uzrokovali su pojavu velikih količina podataka.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Živko Bojović, vanr. prof.

Primenom tradicionalnih baza podataka nije bilo moguće razviti napredne aplikacije i zbog toga se uvode novi pristupi za skladištenje, koji omogućavaju brzu pretragu i analizu velikih količina podataka u realnom vremenu, koji su zasnovani na različitim modelima primene Big Data tehnologije [1]. Termin Big Data (“veliki podaci”) nastao je krajem prošlog i početkom ovog veka. Tada se uglavnom koristio, da bi se u nauci i praksi ukazalo na značaj obrade velike količine podataka, a za šta se smatralo da je moguće uraditi samo korišćenjem superračunara [2].

Na osnovu životnog ciklusa, Big Data se definiše pomoću 5 dimenzija: Volumen ili obim podataka (*Volume*), Raznovrsnost podataka (*Variety*), Brzina generisanja podataka (*Velocity*), uključujući dve nove dimenzije: *Vrednost (Value)* i *Tačnost (Veracity)* [3].

Primena Big data tehnologije ima za cilj da se omogući što kvalitetnije upravljanje količinom podataka koja raste iz dana u dan, velikom brzinom, kako bi se: smanjili troškovi poslovanja, redukovalo vreme koje je potrebno da se izvrše zadaci na računarima i omogućila pojava novih proizvoda i usluga [4].

3. HADOOP PLATFORMA

Big Data obuhvata projektovanje i realizaciju infrastrukture i servisa za skladištenje velike količine podataka, kao i za njihovu pretragu, deljenje, analizu i vizuelizaciju. Jedno od softverskih rešenja za infrastrukturu Big Data je *Hadoop* [4].

Hadoop predstavlja softver otvorenog koda, koristi se za skladištenje i procesiranje velike količine podataka [1]. Razvijen je od strane *Apache* fondacije, a 2003. počinje negov životni ciklus sa projektom „Nutch“ pokrenut od strane Doug Cutting-a koji se odnosio na mogućnost da se vrše milijarde pretraga, kao i obeležavaju milioni web stranica [5]. *Hadoop* omogućava da se veći problemi svedu na manje, da se analiza brzo i jednostavno sprovede, kao i da kompanije lako upravljaju sa velikim količinama podataka [6]. *Hadoop* se sastoji iz četiri osnovna elementa: *Hadoop Common*, *Hadoop Distributed File System (HDFS)*, *Hadoop MapReduce* i *Hadoop YARN* [1].

Glavni cilj HDFS-a je da otkrije greške i automatski ih ukloni, i spreči da pojava greške izazove prekid u radu ili dugo čekanje. HDFS ima *master/slave* arhitekturu, koja je oblikovana prema modelu „rob – gospodar“. Sastoji se od dva čvora: Glavni – master (eng. *NameNode*) i Podređeni čvor - slave (eng. *DataNode*) [5].

HDFS podržava rad sa velikim fajlovima [4]. Iako omogućava pouzdano, dostupno i povoljno skladištenje

podataka, nije dovoljan sam da bi se ostvario odgovarajući nivo adaptacije Hadoop ekosistema. Da bi se taj problem rešio, Google je uveo novi model programiranja, a to je *MapReduce* [7].

MapReduce je softverski okvir, odnosno algoritam koji se koristi za obradu ogromnih količina podataka. Za njega se kaže da je jedno od najefikasnijih rešenja za velike podatke i pruža značajna poboljšanja u velikim aplikacijama koje zahtevaju velike količine podataka u klasterima [11]. Fleksibilan, pruža podršku za rad sa podacima iz različitih izvora (npr. *e-mail*, društveni mediji) koje obrađuje jednostavnim parovima ključ/vrednost. Jedna od glavnih prednosti *MapReduce* programskog modela, je u stvari paralelna obrada podataka, gde se zadaci dele na manje celine koje se paralelno izvršavaju [4].

MapReduce je programski model koji se sastoji od dve celine: *Mapper* - uzima i konvertuje set ulaznih podataka u drugi skup podataka, tako što se pojedinačni elementi raščlanjuju u parove ključ/vrednost i *Reducer* - koristi izlaz sa *Mapper* celine kao svoj ulaz i po određenom kriterijumu kombinuje te podatke da bi dobio manje parove ključ/vrednost [4].

4. DRUŠTVENE (SOCIJALNE) MREŽE

Pojava i razvoj Interneta, pretraživača, kao i društvenih mreža doveli su do naglog porasta količine podataka. Tehnologije kao što su računarstvo u oblaku (eng. Cloud computing) i edge computing imaju ulogu u upravljanju velikim podacima. Takođe, ove tehnologije doprinele su razvoju društvenih mreža. Sa pojavom društvenih (socijalnih) mreža podaci se često dupliraju i njihova obrada postaje još veći izazov [9].

Društvene mreže predstavljaju virtualne zajednice u okviru kojih upotrebom informaciono-komunikacionih tehnologija zasnovanih na Internet aplikacijama ljudi danas razmenjuju informacije, mišljenja, stavove i iskustva [3]. Sadržaj na društvenim mrežama kreiran je od strane korisnika prema njihovim potrebama. Ideja društvenih mreža nastala je još 1890-ih godina od strane Emil Dirkema i Ferdinand Tenis-a, u njihovim teorijama i istraživanju društvenih grupa [10]. Društvena mreža se sastoji od dve celine: čvorišta (eng. *node*) i veze (eng. *edge* ili *links*) [11]. Društvene mreže se mogu opisati pomoću sedam blokova: identitet, konverzacija, deljenje, prisustvo, veze, reputacija i grupe [3]. Društvene mreže omogućavaju korisnicima da kreiraju javne ili polu-javne profile, pregledaju spisak korisnika sa kojima imaju veze, sačuvaju pregled i prenos njihovog spiska veza, kao i od drugih korisnika unutar sistema [10].

Društveni mediji se mogu podeliti u sledeće grupe: *Twitter* - predstavlja jednu od najpopularnijih društvenih mreža koja povezuje uticajne ljude, političare, menadžere i poslovne ljude, *Facebook* - je najpopularnija globalna društvena mreža koju koriste preko milijardu ljudi iz svih strana sveta sa različitim profilima. Pruža razne mogućnosti, kao što su: profil, chat, grupa, događaji, fan stranica, plaćeni oglasi, *YouTube* - predstavlja najveću i najpopularniju društvenu web stranicu u svetu, koja ima ulogu u deljenju video snimaka, *LinkedIn* - je poslovna društvena mreža koja ima ulogu u povezivanju

poslodovaca i ljudi koji traže posao. Sve navedene društvene mreže su i edukativnog karaktera [3].

5. GENERALNA PRIMENA BIG DATA TEHNOLOGIJE

U oblasti informaciono-komunikacionih tehnologija, u najvećoj meri se primenjuje Big Data analitika. Big Data analitika ima potencijal da unapredi poslovanje, kao i svakodnevni život. Mogućnost primene Big Data tehnologije je ogromna, primenjuje se i u drugim oblastima sem informaciono-komunikacionih tehnologija, kao što su: telekomunikacije, finansije, bezbednost, marketing, zdravstvo, sport [12].

5.1. Primena Big Data tehnologije u telekomunikacijama

Big Data tehnologija ima veliku i značajnu primenu u oblasti telekomunikacija. Najznačajni deo telekomunikacija jesu mobilne telekomunikacije u okviru kojih imamo da operateri generišu ogromne količine podataka koje se odnose na telefonske razgovore, SMS i MMS poruke, korišćenje mobilnog Interneta, lokaciju korisnika mobilnog telefona, kao njegovo kretanje, interakciju između operatera i korišćenje mobilnih telefona u inostranstvu [12].

5.2. Primena Big Data tehnologije u finansijama

Primena Big Data tehnologije u finansijama ima značajnu ulogu. Omogućava da se podaci o transakcijama obrade u najkraćem roku i da se na osnovu njih vizuelizuju dobijeni rezultati koji mogu da se koriste za unapređenje poslova [12].

5.3. Primena Big Data tehnologije u bezbednosti

Najvažnija primena Big Data tehnologije jeste u bezbednosti, koja će imati veliki značaj u budućnosti. Big Data tehnologija omogućava da se preventivni rad podigne na najviši nivo i da se pravovremeno identifikuju bezbednosne pretnje ili teroristički napadi.

Kod preduzeća čije poslovanje uključuje potraživanje ili obradu transakcija, Big Data ima najupečatljiviju ulogu primene u otkrivanju prevare [12].

5.4. Primena Big Data tehnologije u marketingu

U marketingu, Big Data predstavlja ostvarenje sna kod ljudi i ukazuje na velika obećanja u toj sferi [10]. Koncept Big Data pomaže ljudima koji se bave marketingom da odgovore na pitanje ko, šta i kada kupuje, dosta brzo, jasno i transparentno, ali zapravo pravi doprinos se ogleda u stvaranju novih oblika vrednosti za kupce. U sferi kupaca, u marketingu, Big Data uključuje ljudske stavove, marketinške kampanje, prodajna mesta, web lokacije, ankete kupaca, društvene medije [13].

5.5. Primena Big Data tehnologije u zdravstvu

Budućnost primene Big Data leži i u oblasti zdravstva (medicina i farmacija). Neki od podataka koji se generišu u oblasti zdravstva su: podaci senzora koji mere vitalne funkcije pacijenata, biometrijski podaci, podaci za ljudske

gene, klinička ispitivanja, za registraciju lekova, laboratorijski, administrativni itd [12]. U medicini, uz pomoć Big Data, možemo unapred otkriti koji pacijenti imaju povećan rizik od oboljevanja, kao što su: srčane bolesti, dijabetes, astma [12]. Takođe, praćenje i predviđanje kretanja epidemije i pojave bolesti [14].

Korišćenjem Big data uočava se mogućnost primene određenih lekova za lečenje nekih bolesti za koje se do tada nije znalo da ih ti lekovi mogu lečiti. Svaki učesnik na farmaceutskom tržištu i apoteke koriste Internet i tako ulaze u digitalni svet. Generišu velike količine podataka, kao što su profili i navike potrošača, kulturološki obrasci, penetracije na tržište, a koji se obrađuju uz pomoć Big Data tehnologije [12].

Big Data tehnologija ima svoju primenu i u brojnim drugim oblastima medicine i farmacije, kao što su unapređenje rada zdravstvenog sistema, zdravstvenog osiguranja, nacionalnih agencija za lekove [12].

5.6. Primena Big Data tehnologije u sportu

Big Data aplikacije se u velikoj meri mogu koristiti da bi se poboljšale sportske performanse u različitim sportovima, kao što su: košarka, fudbal, tenis, bejzbol, formula. Njena primena je dobila značajnu ulogu kod profesionalnog sporta, upravo sa pojavom senzora u loptama, pametnih narukvica, koje prate zdravstveno stanje sportiste, kamera na terenu, GPS uređaja u odeći. Data tehnologija je izvršila analizu velikih serija podataka koji se koriste na sportskim takmičenjima i vrše predviđanje rezultata određenih sportskih manifestacija [12].

Pored navedenih oblasti, moguća primena Big Data je i u predviđanju prirodnih katastrofa i pronalazenju naftnih bušotina na osnovu analize velikih serija geoloških podataka [12].

6. PRIMENA HADOOP-A I BIG DATA U DRUŠTVENIM MREŽAMA

Primena Big Data tehnologije u društvenim mrežama je počela pre oko deset godina i poklapa se sa intenzivnim razvojem Google i Yahoo pretraživača i njihovom odlukom da započnu razvoj softverskih rešenja kojima bi se iskoristile velike količine generisanih podataka i iz njih izvukla određena vrednost. Da bi se ti veliki podaci uspešno, ali optimalno uskladištili, koristi se Hadoop tehnologija blisko povezana sa Big Data konceptom [15].

Društvene mreže su jedan od najrelevantnijih i najrelevantnijih izvora podataka za velike podatke (eng. Big Data). Podaci o društvenim mrežama se dobijaju iz različitih Internet aplikacija i web lokacija (Facebook, Twitter, LinkedIn, Youtube, Google) čiji razvoj stvara novu generaciju ljudi, odnosno omogućava povezivanje korisnika koji međusobno saraduju, dele podatke i komuniciraju [8]. Neke kompanije, kao što su Facebook, LinkedIn, Amazon razvile su svoje platforme preuzimajući neke delove Hadoop-a, i zadržale isti princip rada, odnosno i dalje se skladište velike količine podataka i koristile ga u edukaciji i produkciji. Hadoop omogućava analizu velikih količina nestruktuiranih podataka, kao i

njihovu brzinu. Ti podaci su video i audio podaci, postovi na društvenim mrežama, slike. Hadoop je postao de fakto standard za Big Data, počeo je da raste i da se menja u skladu sa tržištem, tako da je mogućnost primene Hadoop-a svuda gde imate mnogo podataka koje ne možete da obradite tradicionalnim alatima [15]. Veza između Big Data tehnologije i društvenih mreža se zasniva na analitici velikih količina podataka koji mogu nastati iz različitih distribuiranih izvora, ali sa akcentom na društvene medije. Veliki podaci društvenih mreža moraju se složiti kako bi se omogućile bolje usluge korisnicima. Prikupljanje, obrada i analiza velikih podataka o društvenim mrežama je težak zadatak jer klasične metode, algoritmi i alati za upravljanje podacima su postali neadekvatni za obradu velikih količina podataka. Potrebno ih je vizualizovati kako bi se korisnicima omogućio bolji kvalitet, iskustva i veća razumljivost, jer se veliki podaci mogu transformisati u različite formate [8]. Postoji mnogo tehnologija (Amazon, Google) koje rade analizu podataka pošto su određene kompanije uložile veliki novac u razvoj softvera [14]. Društvena mreža Twitter koristi Hadoop distribuiranu bazu podataka – Cassandra [15].

Kao što je već napomenuto, Big Data se obično karakteriše sa „tri V“, odnosno zapremina, brzina i raznovrsnost [10]. Dimenzija *obim* se odnosi na ogromne količine podataka koje pruža društvena mreža. U slučaju da kompanija ne prati konstantno razgovore i strimove na Twitter-u može da propusti pominjanje svog brenda. Big Data može da reši ovaj problem, tako što omogućava da se sačuvaju sva pominjanja koje se posle mogu da pregledaju i analiziraju. Što se tiče dimenzija *brzine*, ona predstavlja brzinu dodavanja podataka. Kod Big Data i društvenih mreža, brzina se odnosi na dodavanje novih informacija iz različitih mreža u bilo kom trenutku, jer društveni mediji nikad „ne spavaju“. Razvojem brzine podataka, povećava se i vreme provedeno na društvenim mrežama. Treća dimenzija *raznovrsnost* se odnosi na različite vrste informacija koje se mogu prikupiti pomoću velikih podataka, tj. na postojanje različitih izvora, kao što su video snimci za nadzor, satelitske slike, mobilni tvitovi, senzori. Raznovrsnost velikih podataka na društvenim mrežama omogućava korisnicima da vide način da se povežu sa željenom publikom [16].

Društvene mreže su postale svakodnevica u svim sferama života, pa i u obrazovanju [14]. Prednost društvenih mreža jeste što značajno unapređuju i povećavaju obim poslovanja. Danas, društvene mreže su zastupljene širom sveta i omogućavaju komunikaciju i deljenje podataka u bilo kom trenutku. Predstavljaju ključ u poslovanju kompanija u budućnosti [10].

7. ZAKLJUČAK ISTRAŽIVANJA STAVOVA GRAĐANA O BIG DATA TEHNOLOGIJI

Istraživanje je sprovedeno za potrebe master rada, imalo je za cilj da se proveru kakvi su zapravo stavovi i znanja ispitanika kada je u pitanju primena aplikacija koje se baziraju na Big Data tehnologiji.

Sprovedeno je tokom juna meseca 2021. god., pomoću online upitnika i obuhvaćeno je 150 ispitanika.

Završetkom istraživanja, dobili su se jasni rezultati o tome koliko su ljudi upoznati sa Big Data tehnologijom. Današnje, moderno društvo je poznato sa savremenim tehnologijama, jer su one deo njihovog svakodnevnog života, tako da je više od 50% ispitanika reklo da je čulo za Big Data tehnologiju, koja će takođe biti deo njihove svakodnevnice, ali manji broj razume pravi koncept Big Data tehnologije i njene uloge.

Kod nas u Srbiji se još uvek ne koristi kao u inostranim zemljama, tako da su odgovori ispitanika bili različiti, nekima pomaže malo, nekima mnogo jer za kratak vremenski period obrade velike količine podataka, ali ispitanici smatraju da Big Data tehnologija ipak ima veliki uticaj u zdravstvu, bezbednosti i naročito na poslovanje kompanija. Ispitanici koriste u većoj meri društvene mreže, najviše Instagram i Facebook, dok YouTube, Twitter i ostale u malo manjoj meri. Skoro 70% ispitanika smatra da Big Data ima ulogu na društvenim mrežama, da omogućava skladištenje velikog broja podataka, brže prikupljanje, analizu i deljenje podataka, otkrivanje prevara, masovnu obradu podataka korisnika društvenih mreža, kao i plasiranje reklama tim istim korisnicima. Najveći broj ispitanika za svoj izvor informacija je odabralo društvene mreže i web stranice, koje imaju veliki uticaj na njihove mišljenje.

8. ZAKLJUČAK

Ogromne količine podataka, koje dolaze iz raznih izvora, neverovatnom brzinom, predstavljale su problem kompanijama u njihovom poslovanju, jer one nisu bile u mogućnosti doneti kvalitetnu i brzu odluku, ali pojavom *Big Data* tehnologije, sve to postaje prošlost. Tada je postalo moguće prikupljati, kombinovati i analizirati u pravcu generisanja informacija u realnom vremenu, i skratiti vreme potrebno za njihovu obradu i analizu, koje su od izuzetnog značaja za donošenje dobrih strateških, taktičkih i operativnih poslovnih odluka, u gotovo svim oblastima poslovanja.

Veliki broj preduzeća u svetu su prihvatili Big Data tehnologiju kao tehnologiju od koje mogu imati velike koristi i deo svojih ljudskih i finansijskih resursa namenili razvoju primene Big Data tehnologije u svom poslovanju. Sa porastom primene ove tehnologije, pojavljuju se i kompanije koje su stručne za njenu implementaciju u svakodnevno poslovanje, kao što su vlada, razne institucije i pojedinci. Pojedine kompanije, velike korporacije, idu korak dalje i osnivaju vlastita odeljenja koja se bave Big Data tehnologijom. Ona pruža velike mogućnosti jer nam pomaže u razvoju novih kreativnih proizvoda i usluga, kao npr. aplikacija na mobilnim telefonima ili proizvodima poslovne inteligencije za kompanije.

Društvene mreže su jedan od najreprezentativnijih i najrelevantnijih izvora podataka za Big Data i njihova veza se zasniva na analizi velikih količina podataka.

Na osnovu svega sprovedenog može se zaključiti da Big Data predstavlja pravu revoluciju poslovanja u mnogim oblastima i da zasigurno menja svet i da će to raditi budućnosti, kao i da možemo iskoristiti bogatstvo podataka koje je do sada samo prolazilo pokraj nas.

9. LITERATURA

- [1]<https://elab.fon.bg.ac.rs/udzbenik-internet-inteligentnih-uredaja/bid-data-i-internet-inteligentnih-uredaja/>
- [2]<http://digitalnapismenost.net/2017/11/23/big-data-drustvena-istraganja/>
- [3]<http://www.ef.uns.ac.rs/predmeti/oas/informacione-tehnologije-i-sistemi/2019-12-19-materijal-za-pripremu-usmenog-ispita-za-pitanja-81-88.pdf>
- [4] Živko Bojović, Jelena Šuh, Petar Bojović, „Softverske tehnologije u računarskim mrežama sa velikim podacima“, Fakultet tehničkih nauka, Univerzitet u Novom Sadu, pp 137.
- [5] Andrea Grozdek, “Tehnologije za rad sa velikim podacima“, Univerzitet u Zagrebu, Prirodno–matematički fakultet, matematički smer, 2017. pp 17-26.
- [6]<https://startit.rs/big-data-hadoop-apache-spark/>
- [7] Aleksandar Simović, Zoran Čirović, “Primena Hadoop okvira u analizi velikih skupova podataka“, INFOTEH-JAHORINA Vol. 15, March 2016, pp 597.
- [8] Gema Bello-Organ, Jason J. Jung, David Camacho, “Social big data: Recent achievements and new challenges“, Computer Science Department, Universidad Autónoma de Madrid, Spain i Department of Computer Engineering, Chung-Ang University, Seoul, Republic of Korea, 2016, pp 45-46.
- [9] Youke Wu , Haiyang Huang, Ningyun Wu , Yue Wang , Md Zakirul Alam Bhuiyan , Tian Wang, “An incentive-based protection and recovery strategy for secure big data in social networks“, 2019, pp 79.
- [10] Slobodan Anđelković, “Savremene društvene mreže“, Univerzitet Singidunum – Departman za postdiplomske studije, master rad, 2015, pp
- [11]<http://www.skladistenje.com/social-network-analysis-analiza-nedrustvenih-mreza-1-dio/>
- [12]<https://www.slideshare.net/IgorLazareviMScEcono/big-data-poslovna-primena>
- [13] Ana Zekavica, Maja Ćirić, Jovanka Vukmirović, “Značaj fenomena Big Data za donošenje marketinških odluka u realnom vremenu“, Ekonomija, stručni rad, 2018, pp 40-41.
- [14] Miloš Popović, “Analiza twitter podataka mogućnosti primene u edukaciji“, Univerzitet Singidunum Beograd, departman za postdiplomske studije i međunarodnu saradnju, doktorska disertacija, Beograd, 2016, pp 18-19.
- [15] Dejan Hadži-Milosavljević, Dušan Starčević, “Hadoop cloud tehnologije“, članak, info M, 2017, Univerzitet u Beogradu, Fakultet organizacionih nauka, pp 4-5.
- [16] Wei Tan, M. Brian Blake and Iman Saleh, Schahram Dustdar, “Social-Network-Sourced Big Data Analytics“, 2013, pp 64.

Kratka biografija:

Ana Petković rođena je u Milićima 1996. god. Master rad na Fakultetu tehničkih nauka iz oblasti Elektrotehnika i računarstva – Telekomunikacioni sistemi odbranila je 2021. god.
Kontakt: ana.petkovic96@hotmail.com