

DDR - ANALIZA TWITTER PODATAKA I MOGUĆNOSTI PRIMENE U EDUKACIJI

By Miloš Popović - II provera

6

UNIVERZITET SINGIDUNUM

BEOGRAD

DEPARTMAN ZA POSLEDIPLOMSKE STUDIJE I

MEĐUNARODNU SARADNJU

**ANALIZA TWITTER PODATAKA I
MOGUĆNOSTI PRIMENE U EDUKACIJI**

- DOKTORSKA DISERTACIJA-

Mentor:

prof. dr Milan Milosavljević

Kandidat:

Miloš Popović, master

BEOGRAD, 2016.

UNIVERSITY OF SINGIDUNUM
BELGRADE
DEPARTMANT OF POSTGRADUATE STUDIES AND
INTERNATIONAL COOPERATION DOCTORAL

TWITTER DATA ANALYSIS AND
POSSIBILITIES OF APPLYING TO
EDUCATION

- DOCTORAL DISSERTATION-

Mentor:

prof. Milan Milosavljević, PhD

Candidate:

Miloš Popović, master

BELGRADE, 2016.

1.	Uvod.....	1
1.1.	Predmet istraživanja	1
1.2.	Hipoteze	2
1.3.	Metode istraživanja	2
1.4.	Ciljevi istraživanja.....	3
1.5.	Struktura rada	3
2.	Primena veb konferencije u e-obrazovanju.....	4
3.	Primena mobilnih telefona u nastavi.....	7
3.1.	Mogućnosti mobilnih telefona	9
3.2.	Operativni sistem Android	10
3.3.	Neki od problema u m-učenju	12
4.	Primena virtuelnih okruženja u nastavi.....	13
4.1.	Second Life	13
4.2.	Open Wonderland	15
5.	Big data	17
5.1.	Definicija Big data	17
5.1.1.	Izvori strukturiranih podataka.....	19
5.1.2.	Izvori nestrukturiranih podataka	21
5.1.3.	Tehnologija Big data	22
5.1.4.	Vizualizacija Big data	27
5.2.	Uticaj Big data na edukaciju	30
5.2.1.	Analitika učenja	32
5.2.2.	Personalizovano i adaptivno učenje.....	33
5.2.3.	Obrazovni podaci	34
5.3.	Primeri primene Big data	34
5.3.1.	Uticaj na nauku	34
5.3.2.	Primena u sportu	35

5.3.3.	Primena u zdravstvu.....	36
5.3.4.	Još neki primeri upotrebe Big data	37
5.4.	Skepticizam	38
11	6. Masovni otvoreni onlajn kursevi (Massive Open Online Courses - MOOC).....	40
6.1.	Pregled poznatih MOOC platformi	41
6.2.	Povezanost MOOC-a i Big data	43
7.	Uticaj i primena društvenih mreža u nastavi.....	44
7.1.	Facebook	49
7.2.	Twitter	50
7.2.1.	Istorijat Twittera i opis podataka	50
7.3.	Upoređivanje Facebooka i Twittera	51
7.4.	Youtube	53
7.5.	Google Plus	54
7.6.	Edmodo	54
7.7.	Wikipedia	55
8.	Big data tehnologije pomoću kojih se analiziraju društveni podaci	56
9.	Spektralna teorija grafova	59
10.	Rezultati istraživanja.....	65
10.1.	Instalacija IBM Infosphere Biginsights tehnologije	66
10.2.	Proces kreiranja fajla za analizu	69
10.3.	Analiza Twitter podataka.....	76
10.4.	Primena podataka u edukaciji.....	82
10.5.	Izrada aplikacije.....	93
10.6.	Upitnik.....	98
11.	Pregled doprinosa disertacije	112
12.	Zaključak.....	113
13.	Literatura.....	114

1. Uvod

Dostupnost i ubrzani razvoj informaciono-komunikacionih tehnologija znatno su uticali na kompletan edukativni proces. Nameću se nova pravila ponašanja u edukaciji kao i nov način kreiranja strategija obrazovanja. Sa ciljem da se poveća efikasnost učenja i podučavanja, neophodno je obrazovni proces obogatiti atraktivnijim sadržajem koji bi studente uključio u proces obrazovanja kao aktivne učesnike. Takođe je potrebno uvideti šta od savremenih tehnologija studenti koriste najviše, i pokušati da se barem na neki način te tehnologije primene u nastavi. U ovom delu rada detaljno će biti opisani predmet istraživanja, polazne hipoteze, naučne metode, ciljevi istraživanja i struktura rada.

1.1. Predmet istraživanja

Informacione tehnologije su postale neophodnost kao i deo svakodnevnice u svim oblastima. Savremeni edukacioni sistemi moraju pratiti neprestani i nezaustavljivi razvoj koji diktiraju nove generacije i njihov način života. Adaptaciju procesa učenja bi trebalo bazirati na dobrim stranama informacionog napretka društvenih mreža: živoj razmeni iskustava i mišljenja, velikim bazama podataka (Big data) i povratnoj informaciji koja je od suštinske važnosti za prenošenje znanja i zadržavanja pažnje i interesovanja sudenata. Ove aspekte bi trebalo iskoristiti na najbolji mogući način i proces edukacije približiti mladim generacijama prikazujući im sadržaj na njima blizak, interaktivan način i istovremenno koristeći potencijale velikih podataka. Brojna istraživanja govore u prilog tvrdnji da je procenat uspešnosti usvajanja i razumevanja gradiva direktno srazmeran učešću informacionih tehnologija i implementaciji novih modela edukacionih sistema.

Visok stepen informatičke pismenosti se danas podrazumeva od najranijeg detinjstva. Obazovanje se više ne bazira na prostom usvajanju pruženih aksioma, već je akcenat na pravilnom izboru, kritičkom mišljenju i pretraživanju velikih baza podataka te dolaženja do rešenja problema na interaktivan način.

Prvi deo ovog rada daje prikaz postojećih rešenja sa aktuelnim informacionim tehnologijama implementiranim u edukaciji danas. S obzirom na to da je 90% svetskih podataka generisano u poslednje dve godine [1], (što jasno pokazuje stepen napretka i brzinu koja je neophodna radi držanja koraka sa novom "net generacijom") drugi deo rada je usmeren ka prikazivanju definicija različitih tehnologija u sistemu obrazovanja. Društvena mreža Twitter pruža korisnicima slobodu kreiranja sadržaja, kao i veliki broj opcija izbora relevantnih informacija. Za modelovanje i

simulaciju društvenih mreža često se koristi opšta teorija grafova, kao i spektralna teorija grafova. Kao alat za analizu društvenih mreža koristi se IBM BigInsights tehnologija koja omogućava vizuelizaciju i analizu velikih podataka. Poslednji deo rada je usmeren ka prikazivanju primera uspešne implementacije navedenih modela kroz kreiranje i upotrebu TDAedu aplikacije [2]. Ova aplikacija objedinjuje navedene pretpostavke i u praksi dokazuje uspešnost kolaboracije standardnih edukacionih sistema sa tehnološkim interaktivnim mrežama. Takođe je kreiran i analiziran upitnik koji pokazuje stav studenata prema uticaju informaciono-komunikacionih tehnologija na poboljšanje kvaliteta nastavnog procesa.

1.2. Hipoteze

U izradi ovog rada polazi se od generalne hipoteze da je novi način za unapređenje performansi obrazovnih sistema implementacija informacija dobijenih analizom podataka sa društvene mreže Twitter, pomoću IBM BigInsights Big data tehnologije.

Izvedene hipoteze:

- Personalizovani pristup u predavanju baziran na velikim podacima dobijenim putem društvenih mreža daje bolje rezultate;
- Kreiranje nastavnog plana koristeći nove tehnologije doprinosi boljem usvajanju gradiva;
- Kritički pristup i statistički uzorak utiču na plasiranje edukativnog materijala;
- Razmena podataka i iskustva sa socijalnih mreža pomaže u rešavanju konkretnih problema u edukaciji.

1.3. Metode istraživanja

U radu su obuhvaćene sledeće naučne metode istraživanja:

- Teoretsko istraživanje sprovedeno je primenom metode sistematske analize domaće i inostrane literature iz domena predmeta istraživanja, sa ciljem da se utvrdi činjenično stanje, definišu problemi i predlože nova rešenja;
- **13** Izvođenje logičkih posledica iz hipoteza i izbor plana i strategije istraživanja;
- Kritička analiza primene Big data u različitim oblastima kao i u edukaciji;
- Analiza prednosti i mana implementacije podataka dobijenih putem društvenih mreža;

- Definisanje i analiza upitnika sprovedenog sa ciljem da prikaže stav studenata o uticaju informacionih tehnologija na kvalitet obrazovanja.

1.4. Ciljevi istraživanja

Opšti cilj istraživanja je podizanje svesti o svakodnevnom korišćenju informaciono-komunikacionih tehnologija u edukativne svrhe, kao i o njihovom uticaju na sam kvalitet nastave. Primarni cilj je modelovanje edukativnog softverskog sistema koji bi omogućio bolji pristup edukaciji i modernizaciji edukativnog sadržaja. Polazne tačke bi bile praćenje informatičkih trendova i interesovanja krajnjih korisnika edukativnog sistema, kao i informacije dobijene analizama podataka sa društvene mreže Twitter putem Big data tehnologije IBM Biginsight.

Na osnovu istraživanja u ovom radu kao i dosadašnjih naučnoistraživačkih iskustava iz ove oblasti, predloženo je i realizovano softversko rešenje za podršku nastavi. Dizajn i implementacija novog softverskog sistema predstavljaju praktični cilj ovog istraživanja.

25

Cilj aplikacije, koja je realizovana u Beogradskoj poslovnoj školi – visokoj školi strukovnih studija, je da se predavačima pomogne da osavremene svoj nastavni sadržaj pomoću podataka koji su trend u svetu, pomoću linkova ka najnovijim blogovima i pomoću komentara najuticajnijih Twitter korisnika.

13

1.5. Struktura rada

Ovaj rad se sastoji od trinaest poglavlja. Posle prvog uvodnog dela je poglavlje u kom se opisuju istorijat i uticaj veb konferencije na edukaciju i drugi faktori koji utiču na e-obrazovanje. Treće i četvrto poglavlje opisuje najzastupljenije modele u savremenoj edukaciji. Detaljno je analiziran i prikazan uticaj mobilnih telefona, a potom i virtuelnih okruženja, na edukaciju. U petom poglavlju je dat uopšten prikaz primene Big data u edukaciji i u drugim sferama života, kao i pregled nekih od tehnologija koje su aktuelne. U šestom poglavlju su predstavljene i analizirane najpoznatije platforme koje nude otvorene onlajn kurseve, koje koriste Big data tehnologiju, u cilju boljeg napretka studenata i osavremenjivanju nastave. U sedmom poglavlju su opisani uticaj i primena društvenih mreža u nastavi. Dat je prikaz najzastupljenijih društvenih mreža i njihovog uticaja, napravljeno je upoređivanje Facebook-a i Twitter-a i obrazloženo je zašto je u ovom istraživanju odluka pala na društvenu mrežu Twiter. Nakon pregleda na koji način informacione tehnologije mogu da utiču na kvalitet edukacije, a posebno sa osvrtom na Big data, realizovan je projekat koji ima za cilj poboljšanje kvaliteta predavanja i testova na nekim predmetima u Beogradskoj poslovnoj školi. U osmom poglavlju se prikazuju Big data tehnologije uz pomoć kojih se analiziraju

društveni podaci. Deveto poglavlje opisuje spektralnu teoriju grafova kao i istraživanja koja su sprovedena na tu temu. U desetom poglavlju opisani su najvažniji rezultati istraživanja. Opisan je način rada tehnologije IBM InfoSphere BigInsights, način instalacije i proces kreiranja fajla za analizu. Objasnjen je sam tok analize Twitter podataka, kako se ti podaci primenjuju u nastavi i opisan je proces izrade aplikacije kao i njena upotreba. U ovom poglavlju se takođe nalazi opis kreiranja upitnika i analiza rezultata ispitivanja stava studenata o uticaju informaciono-komunikacionih tehnologija na nastavni proces. U jedanaestom poglavlju je dat pregled doprinosa disertacije sa objavljenim radovima. Završna poglavlja predstavljaju zaključak svega opisanog, navodi se literatura, indeks slika i tabela korišćenih u radu.

2. PRIMENA VEB KONFERENCIJE U E-OBRAZOVANJU

U ranim edukativnim sistemima na daljinu nastava se uglavnom bazirala na tv i radio obrazovne programe kao i na primenu štampanih materijala. Ovakav vid edukacije je predstavljao odsustvo direktne interakcije između studenta i profesora. Kao dodatak postojećim obrazovnim tehnologijama uvedena je audio konferencija koja predstavlja prvu tehnologiju u obrazovanju na daljinu koja omogućuje dvosmernu komunikaciju između studenata i profesora.

Audio konferencija je omogućila studentima da iz svog doma mogu neometano da uče i da komunikaciju sa profesorima odvijaju u realnom vremenu. U početku je takav vid komunikacije bio na nivou grupe, što je dodatno doprinelo poboljšanju kvaliteta nastave. Nakon audio konferencije, kao naprednija tehnologija, zaživela je i video konferencija koja je pružila novu mogućnost studentima da u realnom vremenu komuniciraju sa profesorima i studentima sa drugih lokacija, preko video tehnologije.

Vremenom je sve veći broj obrazovnih institucija započeo organizaciju i implementaciju nastavnog programa baziranih na video konferencijskim tehnologijama. Ali ovakav vid nastave nije nudio fleksibilnost kakvoj se težilo, jer u ovakvom vidu nastave student je bio u obavezi da fizički bude prisutan na lokaciji na kojoj se organizuju pomenuti nastavni programi. Moralo je da dođe do promene i pojavila se nova tehnologija u polju nastave na daljinu, a to je veb konferencija. Veb konferencija je na efikasan i siguran način uspela da reši postojeće probleme u domenu fleksibilnosti rada.

Jedan savremen obrazovni sistem koristi veb konferenciju u komunikaciji između studenata i profesora. Ukoliko su profesori povezani sa studentima putem širokopojasne mreže, omogućeno je neometano praćenje nastave uz odličan kvalitet videa. Moguće je lako i brzo deljenje podataka u različitim formatima kao i chat između studenata i profesora. Jedan od problema veb konferencije je trošak vezan za licenciranje softvera, kao i još uvek ne tako brza internet konekcija u Srbiji.

U savremenim edukacionim sistemima postoji nekoliko modela konferencije:

- Model jedan na jedan predstavlja komunikaciju između profesora i određenog studenta u individualnoj osnovi.

4

- Model jedan profesor prema grupi udaljenih studenata se bazira na komunikaciji između profesora i grupe studenata na udaljenoj lokaciji. Uglavnom se koristi video konferencija i ovaj model nije mnogo zahtevan.

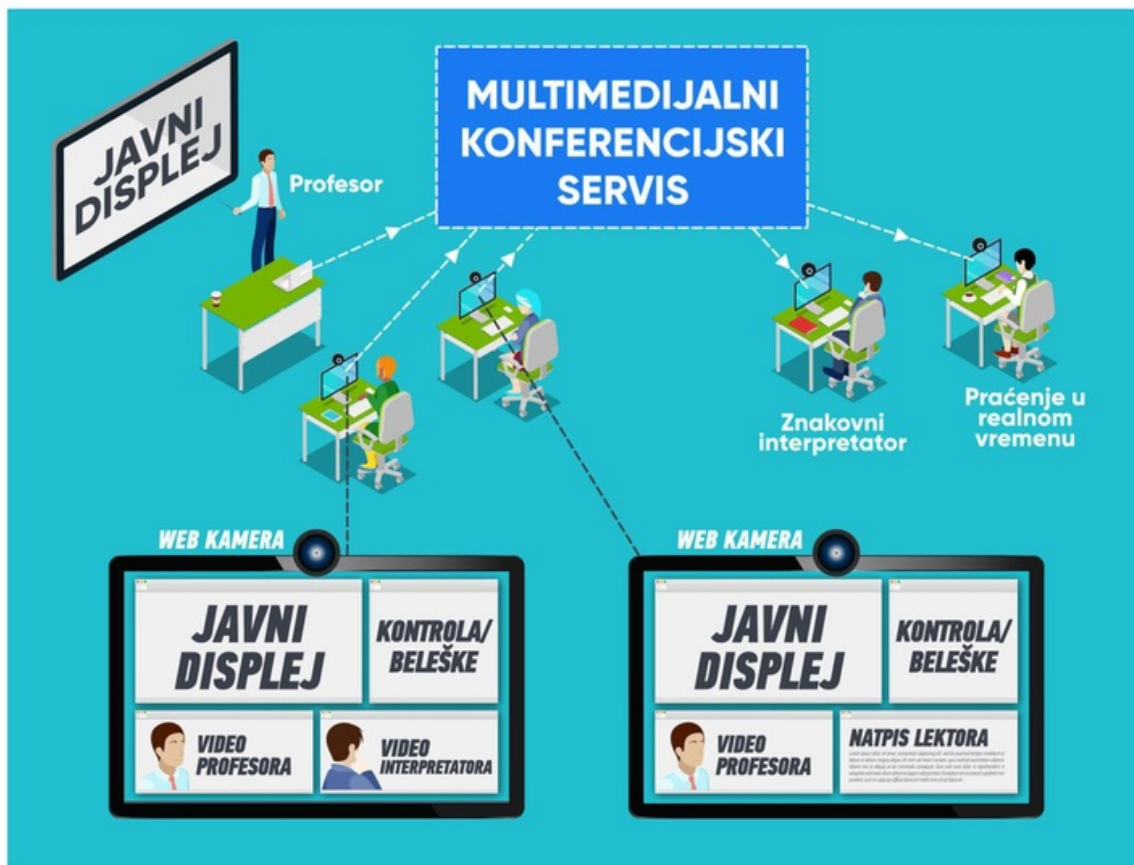
- Model profesor ka udaljenim pojedincima predstavlja komunikaciju između udaljenih studenata koji pojedinačno komuniciraju sa profesorom. Ovaj oblik konferencije nije toliko tehnički zahtevan i verovatnoća za pojavu problema u komunikaciji je znatno smanjena ukoliko broj pojedinaca nije veliki [3]. Primenu ovog modela jedino može da ograniči brzina internet konekcije.

- Model profesor prema grupama studenata Model professor prema grupama studenata se koristi za direktno predstavljanje nastavnog sadržaja studentima koji su na udaljenoj lokaciji. Može se sprovoditi i u audio i u video konferenciji, i tu profesor daje direktne instrukcije vezane za predmet. Može biti odlična dopuna postojećim štampanim materijalima i multimedijalnom sadržaju koji se na primer nalazi na oficijelnom veb sajtu škole ili fakulteta.

4

- Model kolaboracije predstavlja mešavinu modela jedan na jedan i modela profesora ka udaljenim pojedincima.

Veb konferencija je u suštini usredsređena na kolaboraciji između studenata i u potpunosti supstituiše nastavne sisteme. Veb konferencija pruža mogućnost studentima koji nisu fizički prisutni na lokaciji obrazovne institucije da normalno i redovno učestvuju, odnosno angažuju se u nastavi [4].



Slika 1. Veb konferencija

Većina profesora se slaže da ⁴ potencijal tehnologije za primenu u obrazovnom procesu raste srazmerno njenoj sposobnosti da realnije podražava komunikaciju licem u lice. S obzirom na to, veb konferencija igra važnu ulogu u primeni u edukaciji. Što se tiče troškova vezanih za realizaciju veb konferencije, teško je proceniti precizno koliki bi oni bili. Troškovi su uglavnom vezani za industrijski dizajn, upravljanje projektom, vreme profesora koje ulaže u projekat itd. Najjeftiniji model je taj gde professor samostalno organizuje i kreira svoju nastavu putem Power Point prezentacija, četa, audio predavanja, onlajn izvora itd. U tom slučaju profesor sam kreira testove za polaganje ispita. Malo složeniji, a samim tim i skuplji način konferencije bi podrazumevao tim stručnjaka koji bi kreirali kompletan nastavni plan i industrijski dizajn. Tu su i troškovi softvera koji bi se koristili za potrebe veb konferencije, kao i troškovi hardvera tj. mreže i servera. Takode, za kompletnu realizaciju treba pridodati i troškove tehničke podrške u okviru samog fakulteta ili škole. Bitno je navesti i to da su neka istraživanja pokazala da u slučajevima kada se lica studenata pojavljuju u video materijalima, može doći do anksioznosti kod tih studenata i to može dovesti do manjeg učešća u ovakvim modelima konferencije.

Na osavremenjivanje obrazovanja utiču mnogi faktori, kao što su profili studenata i profesora, tehnološki uslovi, akademski faktori itd. Kako se povećava broj studenata koji žele da se obrazuju i studiraju fakulteti sa svojim kapacitetima neće moći da opsluže toliko studenata, jedino rešenje je da se aktivira e-obrazovanje. Fakulteti su prinuđeni da ponude studentima neograničenu mobilnost u učenju. Dolazi do oštre konkurencije između fakulteta i škola i uspeh jedino može biti zagarantovan ako se efikasno koriste novi mediji tj. informacione tehnologije, ako postoji dobar marketing i kontrola kvaliteta. Faktor koji veoma utiče na oblikovanje edukacionih tendencija se odnosi na usavršavanje zaposlenih u obrazovnoj ustanovi. Prioritet je da profesor bude dobro organizovan, "isplaniran" i uvek dostupan studentima jer je komunikacija direktno sa predavačem vrlo izražena u e-obrazovanju. Takođe je potrebno organizovati odgovarajuću podršku i obuku za predavače, radi implemtiranja novih tehnoloških dostignuća u edukacioni sistem. Tu može da se javi problem jer dosta starijih predavača pruža otpor ka svemu tome, takođe se neki od predavača osećaju usamljeno dok rade u virtualnom okruženju jer su fizički odvojeni od studenata. To može itekako uticati na njihovu motivaciju i želju za daljim učestvovanjem u takvim programima.

Jedan od najznačajnijih faktora koji direktno utiče na buduće trendove u e-obrazovanju je tehnološki faktor. Iz dana u dan tehnologija sve više napreduje i samim tim utiče na promenu i budući razvoj edukacije. Širokopojasni internet je sve dostupniji fakultetima i školama, računarska tehnologija je sve jeftinija a sa boljim performansama, video konferencije su vrlo efektivne tako da se polako rešava jedan od najznačajnijih problema koji je najviše opterećivao dosadašnje programe e-obrazovanja. Broj korisnika interneta rapidno raste, a sa tim i kompjuterska pismenost kod studenata i to vrlo povoljno utiče na poboljšanje savremenih edukacionih sistema.

3. PRIMENA MOBILNIH TELEFONA U NASTAVI

Može se reći da m-učenje potiče još od 1901. godine kada je *Linguaphone* krenuo sa časovima učenja jezika preko fonografskih cilindara. 1960. godine počinje upotreba teleprintera koji su korišćeni u edukaciji tako što korisnik odgovara na pitanja koja su prikazana kao alfanumerički znaci od 24 linije sa po 80 karaktera. 70-ih godina se pojavljuju prvi personalni računari i pored teksta počinje da se pojavljuje i grafika. Već 90-ih se pominje m-učenje kao tehnologija, dolazi do razvoja interneta i povezivanja mobilnih kompjutera sa bežičnim mrežama. Pojavom prvih mobilnih telefona počeli su da se razvijaju istraživački projekti koji su se ubrzo počeli koristiti i na fakultetima kao i na radnim mestima.

Mobilni uradaji se u društvu nameću kao idealan instrument mobilnog učenja jer su dostupni većini svetske populacije, a naročito su omiljeni među mladima. Oni su po svojoj prirodi intimni mediji: nose se u džepu i drže se u ruci. Između korisnika i uređaja javlja se bliskost i povećavaju se emocionalni naponi pri njihovom korišćenju. Studenti vrlo brzo postaju privrženi svojim prenosnim računarima. Mobilno obrazovanje nudi puno više pogodnosti nego tradicionalni način školovanja [5].

Pod pojmom mobilnog učenja se podrazumeva oblik učenja na bilo kom mesu i u bilo koje vreme uz pomoć mobilnih uređaja. Telefon pre svega mora da omogući interakciju između studenata i profesora. Izloženost tehnologiji počinje još u ranim godinama, mala deca provode u proseku 2 sata dnevno za televizorom, računarom, tabletom, mobilnim telefonom... Tehnologija nije budućnost, već sadašnjost, to dokazuju učenici u osnovnim školama koji u torbama s knjigama i sveskama uvek pronalaze mesto za svoj mobilni telefon. Napredak u tehnologiji, Wi-Fi, Bluetooth i globalne bežične tehnologije kao GPS, GSM, GPRS, 3G, 4G satelitski sistemi pružaju nove šanse korisnicima. Korišćenjem ovih tehnologija zajedno sa mobilnim uređajima nastala je oblast m-učenja. Ono što je karakteristično za m-učenje jeste nezavisnost od mesta i prilagodljivost načinu života u urbanim sredinama gde se broj korisnika svakim danom sve više povećava. Sama pogodnost m-učenja svodi se na to da je dostupno gde god se student nalazio kao i to što pruža razmenu sadržaja i kolaboraciju među korisnicima. M-učenje takođe omogućava kako onome ko uči, tako i onom ko podučava, da učenje bude u toku i na licu mesta korišćenjem ovih tehnologija, prvenstveno smartphone-a. Time ovi uređaji mogu da postavljaju i ekstraktuju informacije gde god i kad god nastane potreba za tim. Korišćenje mobilnih alata za kreiranje materijala kao i za pomoć pri učenju sada je vrlo važan aspekt kod m-learninga [6].

3.1. Mogućnosti mobilnih telefona



Slika 2. Primeri mobilnog obrazovanja

(Izvor: [https://web2tools-technologies.wikispaces.com/M-u%C4%8Denje+\(mobile+learning\)\)](https://web2tools-technologies.wikispaces.com/M-u%C4%8Denje+(mobile+learning))))

Mobilni telefoni mogu da pomognu u edukaciji na više načina. Osnovne mogućnosti koje pružaju mobilni telefon su glas, sms poruke, zvuk, ekrani sa visokom rezolucijom, velika memorija, veb pretraživači, kamera, android i apple aplikacije itd.

Preko sms servisa se mogu ispitivati mišljenja studenata, mogu se slati razni kvizovi ili testovi pravopisa, matematike ili tome slično. Za učenje stranog jezika i za konverzaciju najviše se koristi zvuk. Postoje softveri za rečnike i knjige kao i mnoštvo aplikacija za testiranje glasovnog engleskog jezika. Skoro svi veb sajtovi su danas dizajnirani specijalno za pametne telefone, a kada student ima mogućnost da preko mobilnog telefona poseti sajt koji mu je potreban onda se može reći da ima sve što mu je potrebno za učenje. Mobilni telefoni mogu i kameru da iskoriste za prikupljanje raznih podataka, pogotovo danas kada su kamere na mobilnom telefonu doslovno zamenile digitalne fotoaparate. Sve navedeno studenti mogu da skladište na svojim telefonima koji se danas uglavnom prave sa 16 ili 32 GB memorije koja je i više nego dovoljna za potrebe studenata. Postoje i skuplji modeli sa po nekoliko stotina gigabajta memorije ali i sa dodatim slotom gde se može staviti memorijska kartica. Android i apple aplikacije poslednjih godina igraju najvažniju ulogu u mučenju. Iz dana u dan se pojavljuju i unapređuju aplikacije iz široke palete oblasti obrazovanja. Od velike koristi je aplikacija Viber, koja je poslednjih godina počela da se koristi na skoro svim pametnim telefonima u svim delovima sveta. Ta aplikacija nudi besplatnu razmenu poruka, glasa

kao i video poziva. Ima mogućnost grupnih poruka i poziva i veoma je pogodna za studente koji žele da se povežu a da ih to pritom ništa ne košta. Dovoljna je samo internet konekcija koja je danas dostupna na skoro svim javnim mestima.

3.2. Operativni sistem Android

Android predstavlja operativni sistem za mobilne uređaje. Razvijen je u kompaniji Google i bazira se na Linux kernelu i GNU softveru. Android platforma za mobilne telefone i tablete je danas najpopularnija na svetu. Uz nju je moguće brzo povezivanje i deljenje informacija, cloud-sync, multi-tasking kao i milioni javno dostupnih Google aplikacija. Android je započet 2003. godine kao operativni sistem za telefone i tablete a u masovnoj upotrebi je tek od 2008. godine. Njegov zadatak je da uspostavi vezu između uređaja i svih ostalih programa i aplikacija. Android telefoni se mogu koristiti za telefoniranje i slanje SMS i MMS poruka. Sa Androidom je lakši pristup internetu, jer ima niz aplikacija koji su specijalizovane za pojedine servise i sajtove.

Android poseduje i GPS sistem. Takođe raspolaže i velikim brojem igrica i aplikacija namenjenim učenju, poslovanju, komponovanju muzike itd. Aplikacije za Android se uglavnom skidaju sa interneta, a najveći i najsigurniji je *Google Play*. Na njemu postoji ogroman broj besplatnih aplikacija. Android je od Linuksa preuzeo filozofiju otvorenog koda. Svakoga dana na hiljade nezavisnih programera unapređuje i proverava programski kod. Za Android je trenutno dostupno više miliona različitih programa, a broj instalacija u svetu prelazi deset milijardi [7].

Android se može koristiti i na lap-top računarima, elektronskim knjigama, takođe ima upotrebu na digitalnim satovima, kamerama, televizorima i još mnogo čemu drugom. To je moguće zbog njegovog otvorenog koda i arhitekture.



Slika 3. Popularan izgled Android operativnog sistema

(Izvor: <http://www.aptech-education.com/courses-smart-professional-android-mobile-app-development.aspx>)

Neke od pogodnosti koje donosi Android, a koje se ujedno mogu iskoristiti u edukaciji:

- Brzo i lako pretraživanje. Pomoću Androida moguće je videti nekoliko stranica odjednom. Moguće je sačuvati bilo koju veb stranicu.
- Upravljanje fotografijama i povezivanje. Uz pomoć određenih aplikacija sa samo jednim klikom moguće je podeliti fotografije. Moguće je koristiti kameru da bi se ostvario video chat preko Google Talk-a, Skype, Vibera, Facebook-a...
- Dostupnost mail-ova. Zahvaljujući Gmail-u vrlo se lako šalju i čitaju e-mailovi preko mobilnih uređaja. Svi vodeći svetski veb mail-ovi imaju svoje aplikacije koje se koriste na Androidu, vrlo lako i pregledno.
- Mape sa navigacijom. Uz Android dolazi i Google Maps aplikacija koja danas u potpunosti zamenjuje GPS navigacije, pogodna je za korisnike jer može da radi i u offline režimu ukoliko se prethodno download-uje određena mapa.

- Razmena podataka je vrlo jednostavna i laka, koristi se Bluetooth ili Near Field Communication (NFC).
- Upravljanje pomoću glasa je novija tehnologija koja se brzo razvija i podržava različite jezike.
- Noviji Android telefoni se mogu otključati pomoću opcije „face unlock” ili pomoću otiska prsta. Nisu potrebne nikakve šifre.
- Prilagodljiv potrebama. Pomoću dodataka (widgets) koji olakšavaju upravljanje aplikacijama sa početnog (Home screen) ekrana. Moguće je dodati News Feed sa različitih društvenih mreža, e-mail poruke, kalendar vremensku prognozu itd.

3.3. Neki od problema u m-učenju

Pri rukovanju mobilnim telefonom javljaju se i pojedini problemi. Nedostaci se odnose često i na kratko trajanje baterije, uglavnom traje najviše 12 sati sa aktivnim korišćenjem. Takođe je problem ne tako dobra konekcija u svim delovima države. Veliki problem predstavlja kontrola informacija kao i cena uređaja koja može da pređe i 1000 evra. Programi m-učenja zahtevaju dodatno obrazovanje kako studenata tako i profesora koji mora da prilagodi svoje edukativne veštine novim tehnologijama.

Što se tiče sigurnosti, Android je podložan virusima i drugim oblicima napada, najveći problem predstavlja ljudski faktor. Najčešće se dešava da se korisnik složi sa bilo čim što program traži, pa se na taj način često dešava da program u sistem prodre mnogo dublje nego što bi smeo. Ljudi slučajno dozvoljavaju nekim programima da prate gde se nalaze, sa kim razgovaraju i razmenjuju poruke, čak i da krađu njihove fotografije iz telefona ili telefonske imenike. Google sa svog marketa uklanja aplikacije koje zloupotrebljavaju poverenje korisnika i mogućnosti sistema, ali uglavnom tek nakon što je šteta učinjena [8].

Moglo bi se zaključiti da, i pored ovih spornih tačaka, učenje preko mobilnih telefona ima veliki potencijal u budućim edukacionim sistemima.

4. PRIMENA VIRTUELNIH OKRUŽENJA U NASTAVI

4.1. Second Life

Second Life predstavlja popularni virtualni trodimenzionalni svet na internetu u kom korisnici prave svoje avatare pomoću kojih mogu da uče, da se zabavljaju ili da rade neke stvari koje možda ne bi radili u realnom svetu. Mogućnosti su ogromne. Second Life je pokrenut u San Francisku 2003. godine kao softverski paket gde ljudi koji koriste platformu istovremeno predstavljaju i "stanovnike" virtuelnog sveta. Broj tih stanovnika koji se nazivaju avatari trenutno iznosi preko 25 miliona [9]. Second Life se nalazi na serveru kompanije *Linden Lab* a programski jezik koji se koristi je *Linden Script Language (LSL)*. U njemu se nalaze predstavništva zemalja, mnoštvo velikih kompanija koje se bave trgovinom, zabavom i promocijama. U tom digitalnom univerzumu živa globalna zajednica prevazilazi granice kulturne, etničke i generacijske prirode. Korisnici imaju neverovatne i neograničene mogućnosti za istraživanje. Moguće je pronaći odlično mesto gde bi se izgradila kuća ili neka poslovna platforma, moguće je putovati, prikazivati, skladištiti virtuelne kreacije, organizovati događaje i poslovne sastanke itd. Second Life je bogat dešavanjima najrazličitijih sadržaja koje plasiraju razne svetske kompanije. Tu se dešavaju razne premijere filmova i koncerata. Organizuju se razna predavanja i treninzi a mogu se sresti i neki od najpoznatijih svetskih univerziteta kao što je Harvard.



Slika 4. Second Life u edukaciji

Neki fakulteti su prihvatili Second Life kao savremeni edukacioni model. On nudi tri načina učenja: dopunsko, celokupno i pomešano. Uz pomoć demonstracija u Second Life-u dopunski oblik učenja pruža studentu dopunske informacije i proširuje osnovni stepen obrazovanja. Oblik celokupnog učenja je podržan od strane fakulteta i omogućava studentu uvid u ceo nastavni plan u virtuelnoj učionici. Pomešano učenje je kombinacija virtuelnog i stvarnog učenja koje poboljšava obrazovanje studenata.

Kada se suočavaju sa „digitalnom generacijom“ studenata predavači bi trebalo da se više zainteresuju za njihove rezultate učenja. Uživajući sve pogodnosti onlajn učenja i prihvatajući inovativnu tehnologiju potrebno je stvoriti interaktivno virtuelno okruženje za predavanje i učenje u Second Life-u. Osnovni razlozi za sve veću primenu Second Life-a u obrazovne svrhe su pristupačnost, niska cena i laka integracija sa ostalim servisima dostupnim uz pomoć interneta.



Slika 5. Prikaz Second Life učionica

Različite konferencije se osnivaju od strane rastućeg broja organizacija, velikih kompanija i univerziteta, organizuju se i posete virtuelnim muzejima, bibliotekama, bioskopima, izložbama i to sve u cilju edukacije. Sve češće se uspešno organizuju razne konferencije u potpuno virtuelnom okruženju sa učesnicima iz celog sveta [10]. Pojedine organizacije i fakulteti su u potpunosti iskoristili rastući potencijal koji pruža Second Life omogućivši studentima da dožive iskustva koja ne bi mogli u stvarnom životu. Američka organizacija za okeane i atmosferu je u svojoj simulaciji omogućila studentima potpuno interaktivnu obrazovnu demonstraciju o okeanima i vremenskim uslovima, kao što je vožnja u podmornici, dve različite demonstracije cunamija, temperaturne mape, let avionom u uslovima uragana i topljenje glečera.



Slika 6. Vežbe u virtuelnoj ambulanti

4.2. Open Wonderland

Open Wonderland predstavlja virtualni svet koji je baziran na programskom jeziku Java, otvorenog je koda i koristi se besplatno. Korisnik koji zna da koristi Javu može da pravi sam svoje okruženje kako njemu odgovara. S obzirom da server dolazi sa veb administratorskim interfejsom omogućena je promena i nadogradnja serverske aplikacije. Administrator može da iskoristi postojeće module za kreiranje okruženja, a može se koristiti i sa sledećim 3D alatima da bi sam kreirao okruženje: *3ds Max, Blender, Lightwave i Maya, Solid Works...*

Opcija koja daje prednost ovom virtualnem svetu u odnosu na druge, u korišćenju u svrhu obrazovanja, je ta što svaki korisnik može sam da kreira 2D desktop aplikacije uz pomoć Office paketa ili PDF-a, avi fajlova, mp3, jpg itd.

U poslednjih nekoliko godina, posebno zbog veće dostupnosti kvalitetnijeg softvera i hardvera, 3D okruženja su postala uobičajena tehnologija. Veliki istraživački izazov za tehnologiju koja se koristi u obrazovanju je da se stvori mogućnost lakšeg učenja ali i bolje komunikacije između studenata i predavača čak i kada su prostorno udaljeni.

Jedan od zanimljivih primera, kako se Open Wonderland virtualni svet može iskoristiti u edukaciji, je taj da su zajedničkim naporima tri univerziteta (Tehnološki univerziteti iz Graca, Perta i

Kembridža iz različitih krajeva sveta napravili prvo zajedničko virtualno okruženje za višestruke eksperimente iz oblasti fizike. Na slici 7. je prikazano kako izgleda elektromagnetno polje u pomenutom virtualnom okruženju.



Slika 7. Open Wonderland platforma [11]

Novembra 2011. godine u Madridu je održan evropski Samit 3D obrazovanja. Predavači su sa 6 različitih lokacija koristili isto virtuelno okruženje da bi predstavili 9 projekata napravljenih u Open Wonderland-u. Na Essex univerzitetu studenti su mogli da prate samit zahvaljujući Majklu Gardneru koji je dešavanja u virtuelnom okruženju prikazivao studentima u učionici. Jedan od učesnika je prvi prezentovao „svet“ koji je kreiran na univerzitetu Karlos III, pod nazivom „Gran Via“ koji se koristi za učenje španskog jezika. Projekat „ISocial Garden World“ napravljen je na univerzitetu Misuri koji pomaže učenje dece kod kojih je dijagnostikovao autizam [12].



Slika 8. Open Wonderland učionica

Kada se uporede Open Wonderland i Second Life može se doći do zaključka da su veoma slične platforme, ali ima nekih stvari koje idu u korist Open Wonderland-a. Može se reći da izgleda kvalitetnije i stabilnije od Second Life-a i što je najbitnije moguće je kreirati sopstveno virtualno okruženje besplatno. Interfejs je lakši za korišćenje ali s obzirom na to da je Open Wonderland novija virtualna zajednica, samim tim i broji manje korisnika. Mana je ta što ne nudi gotove sadržaje kao što to radi Second Life.

5. BIG DATA

5.1. Definicija Big data

Svaki put pri obavljanju radnji poput kucanja, klika, dodira tač skrina - drugim rečima, stupanja u kontakt sa digitalnim uređajima podaci se prikupljaju, i čuvaju. Upotrebom se kontinuirano ostavlja trag koji se prati, analizira i kojim se čak trguje. Negde se smisleni prikaz formira što zauzvrat omogućava uređaju upoznavanje korisnika lično. Ovakav odnos upoznavanja sa digitalnim uređajem, preko društvenih mreža, pretraživača, sajtova za kupovinu i mobilnih aplikacija, daje osećaj efikasnosti bez koga život postaje nemoguć. Gedžeti snimaju koliko struje svaki uređaj u kući utroši, potrošačka genomika generiše personalizovane lekove, "Nike fuel" narukvica prati lične podatke u toku vežbanja, GPS pomaže povratku kući sa bilo kog mesta, Facebook vesti biraju za prikaz poruke najbližih prijatelja, Google olakšava pretrage izborom relevantnih stvari itd. Lični digitalizovani podaci i način kako su oni predstavljeni postaju nova dimenzija onoga što čini našu "stvarnost". Da bi ovo bilo moguće, potrebno je stalno snabdevati, uglavnom pasivno, more podataka koji se pretvaraju u smislenu informaciju. Ova pojava je ono što obuhvata medijsku frazu Big data [13]. Ogromni porast aluzija na Big data kroz medije potvrđuje neosporni uticaj na sve

oblasti društva, od posla do obrazovanja. Big data je termin koji opisuje skladištenje i analizu velikih i složenih skupova podataka koristeći niz tehnika.

Big data predstavlja mogućnost upravljanja velikim količinama različitih podataka razumnom brzinom i u odgovarajućem vremenskom okviru da bi se omogućila analiza tih podataka u realnom vremenu. Prikupljanje podataka se vrši iz raznih drugačijih izvora i nalaze se u drugačijim oblicima. Eksplozijom razvoja senzora, pametnih uređaja i socijalnih mreža podaci su postali složeni prvenstveno zato što sada ne uključuju samo tradicionalne strukturirane podatke, već i nestrukturirane ili polustrukturirane podatke.

Strukturirani podaci opisuju podatke koji su grupisani u relacione šeme (redovi i kolone u okviru standardnih baza podataka). Organizacija ovih podataka daje mogućnost izvršavanja jednostavnih upita koji mogu vratiti korisne informacije za poslovanje.

Polustrukturirani podaci predstavljaju podatke koji su grupisani u neku fiksiranu šemu. Ovi podaci su skoro uvek teško razdvojivi ali ujedno i sadrže neke od oznaka koje mogu da pomognu u hijerarhijskoj organizaciji ovakvih tipova podataka.

Nestrukturirani podaci predstavljaju uglavnom slike i video ili audio fajlove. Najčešće ih je veoma teško ubaciti u relacione tabele baza podataka radi analize ili izvršavanja upita nad njima.

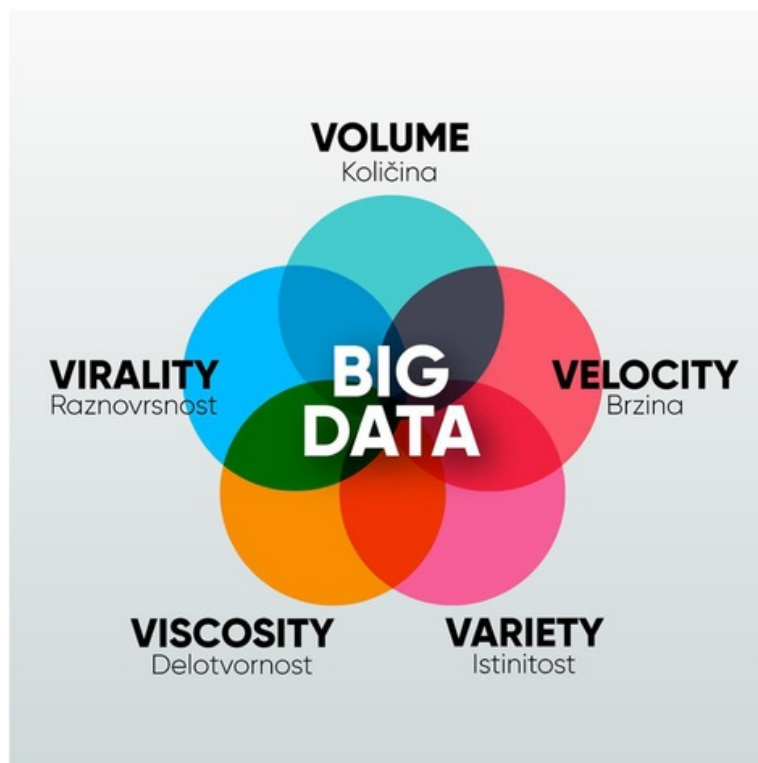
Neki od primera generisanja velikih količina podataka:

- Sistemi radio frekvencija generišu 1000 puta više podataka od tradicionalnih bar kod sistema.
- 10 000 transakcija plaćanja kreditnom karticom se obavi svake sekunde u svetu.
- Korisničke transakcije, više od 1 000 000 u satu obrađuje *Walmart*.
- 340 miliona tvitova se pošalje dnevno. To je približno 4 000 tvitova u sekundi.
- Facebook ima više od 1,6 milijardi aktivnih korisnika koji svakodnevno generišu podatke svojom međusobnom interakcijom.
- Više od 5 milijardi ljudi koristi mobilni telefon kako u razgovoru, tako i u slanju poruka, surfovanju internetom, tvitovanju itd [14].

Big data tehnologija radi kroz masivne količine informacija u realnom vremenu i ranije prikupljenih informacija, kako bi pronašli neviđene obrasce i otkrili naoko besmislene stvari koje mogu da

dovedu do novih saznanja, i ukažu na mogućnosti za nove usluge i proizvode. Štaviše, omogućava se razvoj načina za efikasnije funkcionisanje, poboljšanje transparentnosti i odgovornosti institucija.

Najnoviji tehnološki razvoj, posebno cloud bazirani serveri i razvoj mreža za prenos podataka, omogućavaju i povećanje potencijala Big data koja ima veliki uticaj na društvo. Svakodnevno rastu mogućnosti koje se nude ljudima kao digitalnim korisnicima, radi jačanja vrednosti informacije.



Slika 9. Jedna od definicija Big data

5.1.1. Izvori strukturiranih podataka

Iako se čini da su strukturirani podaci dobro poznati, zapravo, strukturirani podaci u svetu Big data pristupa dobijaju novu ulogu. Razvoj tehnologije omogućava pojavu novih izvora strukturiranih podataka, često u realnom vremenu i u velikim količinama.

Izvori podataka se dele u dve kategorije. U prvu kategoriju spadaju ljudski generisani podaci, i kao što sama reč kaže njih obezbeđuju ljudi u interakciji sa kompjuterima. U drugu kategoriju spadaju mašinski generisani podaci i njih proizvode računari tj. mašine bez uticaja ljudi.

Mašinski generisani strukturirani podaci mogu da uključuju:

Senzorske podatke - Primeri uključuju pametne merače (npr. elektronska brojila za merenje potrošnje električne energije), radio frekvencijske identifikacije (RFID) oznake, podatke medicinskih uređaja, GPS podatke. Tehnologija koja je sve popularnije je RFID tehnologija gde se koriste minijturni računarski čipovi da bi se uređaji pratili sa udaljenosti. Kada prijemnik dobije informacije one mogu biti prosledene serveru gde će biti analizirane. Mnoge kompanije su zainteresovane za ovu tehnologiju zbog pristupačne cene i lakog upravljanja transportom robe, kao i kontrolu inventara. Još jedan primer izvora senzornih podataka su pametni telefoni koji imaju senzore kao što je GPS koji mogu biti korišćeni za razumevanje ponašanja potrošača na novi način.

Veb log podatke - Kada serveri, aplikacije, mreže i slično, rade oni beleže različite podatke o svojoj aktivnosti. Vrlo često količina ovih podataka postane ogromna, a ovi podaci mogu biti od koristi za, na primer, predviđanje narušavanja bezbednosti.

Podatke u trenutku prodaje - Kada radnik na kasi očita bar kod bilo kog proizvoda koji se kupuje, generišu se svi podaci vezani za proizvod. S obzirom da ljudi svakodnevno kupuju različite proizvode, količina podataka koji se prikupe ovom metodom je ogromna.

Finansijske podatke - Dosta finansijskih sistema je danas programirano, njihov rad se zasniva na predefinisanim skupovima pravila što automatizuje proces. Kao najbolji primer može se uzeti berza i podaci o trgovanju. Ti podaci su strukturirani i sadrže npr. oznaku kompanije i vrednost u dolarima. Neki od ovih podataka su mašinski generisani a neki ljudski generisani.

Primeri ljudski generisanih strukturiranih podataka mogu da uključuju:

Ulazne podatke - Ovo je bilo koji tip podataka koji čovek može uneti u računar, kao što je ime, prezime, godine starosti, prihod, odgovori na ankete i slično. Ovi podaci mogu biti korišćeni za razumevanje osnovnog ponašanja potrošača.

Klik podatke - Svaki put kada se klikne na link na sajtu podaci se generišu. Ovi podaci mogu biti analizirani da bi se odredilo ponašanje potrošača i obrasci kupovine.

Podatke vezane za igre - Svaki potez koji se napravi u igri može biti zabeležen. Ovi podaci mogu biti korisni za razumevanje kako krajnji korisnici igraju igru.



Slika 10. Svakog minuta se generišu ogromni podaci [15]

5.1.2. Izvori nestrukturiranih podataka

Nestrukturirani podaci su podaci koji ne prate neki definisani format. Ako je 20% podataka koji su dostupni preduzećima strukturirano, preostalih 80% je nestrukturirano. Nestrukturirani podaci su zapravo podaci koji se najčešće sreću. Do skoro tehnologija nije podržavala druge načine rada sa ovim podacima osim skladištenja i ručne obrade. Nestrukturirani podaci se mogu naći svuda, većina ljudi i organizacija funkcioniše na osnovu nestrukturiranih podataka. Kao i u slučaju strukturiranih podataka i nestrukturirani podaci mogu biti mašinski ili ljudski generisani. Neki primeri mašinski generisanih nestrukturiranih podataka su:

Satelitske slike – Prilikom satelitskog nadgledanja zemlje, vlade prikupljaju razne podatke. U ovu vrstu podataka spadaju i podaci o vremenskim prilikama. Na primer, GoogleEarth poseduje ogromnu količinu satelitskih snimaka koje obrađuje i spaja na odgovarajući način.

Naučni podaci - Ovo uključuje seizmičke slike, atmosferske podatke, fiziku visokih energija, itd.



Grafikon 1. Procena odnosa sačuvanih strukturiranih i nestrukturiranih podataka [16]

5.1.3. Tehnologija Big data

Veliki broj novih tehnoloških dostignuća omogućava organizacijama da iskoriste veliku količinu podataka kao i da ih efikasno analiziraju. Neke od karakteristika su:

- Jeftino i veliko skladište za podatke, uz mogućnost serverske obrade.
- Brži procesori.
- Dostupne mogućnosti za veliku memoriju, kao što je Hadoop.
- Nove tehnologije vezane za skladištenje i obradu podataka, namenjene baš za velike i obimne podatke, uključujući i nestrukturirane podatke.
- Paralelna obrada, klasterovanje, MPP, virtualizacija, velika grid okruženja, visok nivo propusnosti i mogućnosti povezivanja
- Rad u cloud sistemu i druga fleksibilna rešenja za rad sa resursima.

Tehnologije koje se svrstavaju pod “Big data” tehnologije ne podržavaju samo mogućnost prikupljanja velike količine podataka, one daju mogućnost za razumevanje tih podataka kao i

izvlačenje nekih vrednosti. Glavni cilj svih organizacija koje imaju pristup kolekcijama velikih podataka trebalo bi da bude to da iskoriste većinu relevantnih podataka u svom poslovanju za donošenje raznih poslovnih odluka.

Pravi proboj u tehnologiji Big data desio se kada su kompanije kao što su “Yahoo!”, “Google”, i “Facebook” došle do saznanja da se pomoću ogromnih podataka koje su generisali njihovi proizvodi, može zaraditi. Ove kompanije su bile pred zadatkom da nađu način u vidu nekih novih tehnologija koje će im omogućiti da čuvaju, pristupaju, obrađuju i analiziraju ogromne količine podataka u realnom vremenu, na takav način da mogu da zarade i na pravi način iskoriste količinu podataka koju poseduju i koji učestvuju u njihovim mrežama [17]. Došli su do nekih rešenja koja trajno menjaju tržište upravljanja podacima. Na primer Hadoop, MapReduce i Big Table predstavljaju novine koje su dovele do neke nove ere upravljanja podacima. Ove tehnologije imaju akcenat na sposobnost obrade velikih količina podataka na efikasan i blagovremen način, na način koji je isplativ i koji ne zahteva velike troškove.



Slika 11. Big data ekosistem

(Izvor: http://www.slideshare.net/jyrki_m/big-datainbigcompanies-iiia)

- MapReduce

MapReduce je rešenje koje je predstavio Google kao način efikasnog izvršavanja skupa funkcija nad ogromnim količinama podataka na serijski način. Komponenta “map” raspoređuje

programerski problem ili zadatak na veliki broj sistema i rukovodi postavljanju zadataka na način koji podrazumeva balansirano opterećenje i upravlja oporavkom od grešaka. Nakon što se završi distribuirana obrada, poziva se druga funkcija nazvana “reduce”, koja spaja sve elemente nazad zajedno, da bi obezbedila rezultat.

MapReduce je programerski model koji obrađuje velike skupove podataka pomoću paralelnih, distribuiranih algoritama u klasteru. Jedan MapReduce program obuhvata Map() proceduru koja vrši filtriranje i sortiranje (na primer sortiranje studenata po imenu u redove, po jedan red za svako ime) i Reduce() procedura koja vrši operaciju agregacije (na primer broj studenata u svakom redu). MapReduce sistem (može se reći i infrastruktura ili frejmwork - okvir) upravlja distribuiranim serverima i uopšteno celim procesom. Sistem izvršava različite zadatke paralelno, upravlja svim komunikacijama kao i prenosom podataka između različitih delova sistema, u isto vreme obezbeđujući sistem od grešaka. MapReduce biblioteke se pišu na raznim programskim jezicima. Sledi tabela koja prikazuje razliku između relacionih baza podataka i MapReduce [18].

Poređenje relacionih baza i MapReduce	RDBMS	MapReduce
Podrška šema	 Striktno šema	Nema šeme (Posle primene HDFS  blokova, MapReduce može da rukuje strukturiranim podacima)
Index	 Dostupni su razni indexi	 Nema indexa
Programiranje Model / elastičnost	SQL, deklarativni jezik, dovoljan  za sumiranje podataka, lak za korišćenje Za složene analitike potrebno je  UDF programiranje	"Map" i "Reduce" funkcija, interfejs  je na niskom nivou, težak za korišćenje  Veoma fleksibilan
Optimizacija	 Složen "optimajzer" pokušava da pronađe optimalno izvršenje planova za upit	 Mora da se poboljša
Izvršna strategija	Resetuje ceo upit kada se desi  greška	Posao je podeljen u zadacima, u slučaju greške resetuje samo taj  zadatak
Tolerancija grešaka	 Niska	 Visoka
Skalabilnost	Najniža, uglavnom skalira gore, ograničeno skaliranje na stotine  čvorova	 Visoka, može se skalirati veliki klaster, više od 1000 čvorova
Raznovrsnost	 Ograničena podrška	 Potpuna podrška

Tabela 1. Razlika između relacionih baza podataka i MapReduce

- Big Table

Big Table je rešenje razvijeno od strane kompanije Google, kao distribuirani sistem za skladištenje podataka koji je predviđen da upravlja veoma skalabilnim strukturiranim podacima. Podaci su organizovani u tabele sa redovima i kolonama. Za razliku od tradicionalnog relacionog modela baze podataka, Big Table predstavlja proređenu, distribuiranu i trajnu sortiranu višedimenzionalnu mapu. Big Table se uglavnom koristi za čuvanje velikih količina podataka na običnim serverima. Big Table mapira dva proizvoljna stringa (ključ koji se odnosi na red i ključ koji se odnosi na kolonu) i vremenski trenutak (trodimenzionalno mapiranje) u neki vezani niz bitova. Big Table je predviđen da može da ide do nivoa petabajta, rad na preko stotinu hiljada mašina koji omogućava jednostavno dodavanje novih mašina u sistem i njihovo momentalno uključanje u rad na način koji ne zahteva ni prekidanje rada sistema niti bilo kakvo ponovno konfigurisanje.

- Hadoop

Među postojećim projektima jedan od najznačajnijih, široko korišćenih, je Apache Hadoop. To je open source softverski projekat koji omogućava distribuiranu obradu velikih setova podataka preko klastera servera. Hadoop olakšava analizu ogromne količine i brzine nestrukturiranih podataka koji trenutno proizvedeni kao video i audio podaci, postovi preko društvenih medija, slika, itd. Hadoop se nalazi u 37% svih open source Big data tehnologija, kao i u 47% komercijalnih Big data tehnologija [19]. IBM, Cloudera, Microsoft, Hortonworks, Amazon Web Services itd. Distribuiraju Hadoop svojim proizvodima.

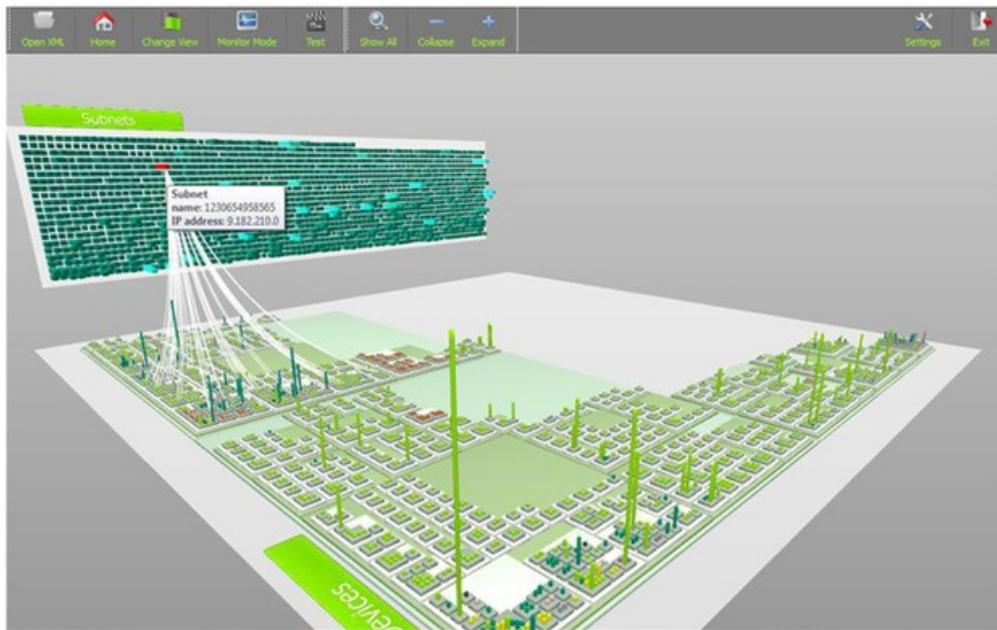
Inovatori sistema za pretraživanje kao što su Yahoo i Google su uvideli da bi trebalo da nađu način kako da izvuku smisao i neku vrednost iz ogromnih količina podataka koje njihovi sistemi prikupljaju. Oni su morali da razumeju u isto vreme koje informacije prikupljaju, kao i kako da povećaju svoje prihode, poboljšaju svoje poslovanje, i da sve te informacije uklope u njega. Hadoop dozvoljava kompanijama da na lak način upravljaju velikim količinama podataka. Hadoop omogućava da veliki problemi budu razbijeni na manje tako da analiza može da se izvrši brzo i jeftino. Razbijanjem tih velikih problema na manje delove koje je posle moguće obrađivati paralelno, i po završetku obrade te informacije se prikupljaju i grupišu radi izdavanja krajnjih rezultata. Hadoop je softverski frejmwork izveden iz MapReduce i BigTable sistema. Hadoop dozvoljava aplikacijama baziranim na MapReduce sistemu da se izvršavaju na velikim klasterima običnog hardvera. Hadoop je dizajniran da paralelizuje obradu podataka koristeći čvorove za povećanje brzine izračunavanja i smanjenje odziva. Hadoop se sastoji od dve glavne komponente, visoko skalabilnog distribuiranog fajn sistema koji podržava i količinu podataka koja se meri u petabajtima, dok je druga komponenta MapReduce sistem [20] [21].

Sve značajnija postaje obrada podataka u realnom vremenu - stream computing. Od tehnologija koje se time bave treba pomenuti IBM Stream Computinga, Microsoft StreamInsight i ORACLE Complex Event Processing.

5.1.4 Vizualizacija Big data

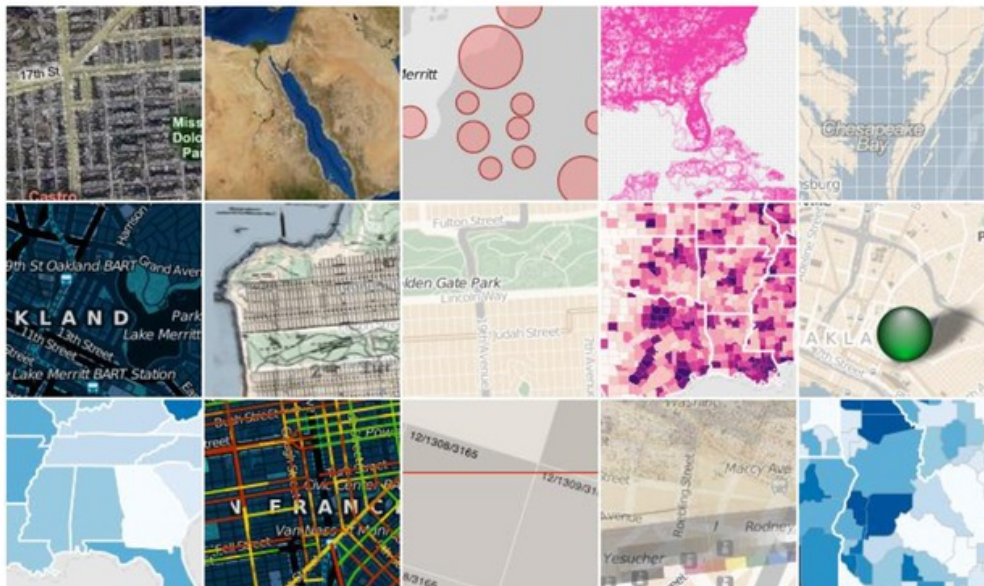
Vizualizacija podataka je predstavljanje podataka u slikovnom ili grafičkom formatu. Vekovima su ljudi zavisili o vizuelnim predstavljanjima podataka, kao što su grafikoni i mape, da bi lakše i brže razumeli informacije. Studije pokazuju da mozak obrađuje slike 60000 puta brže nego tekst. Postoji puno softvera koji omogućavaju korisnicima da vide analitičke rezultate prezentovane vizuelno, prikazuju relevantnost među milionima promenljivih, pa čak i da predviđaju buduće događaje. Vizuelizacije podataka su svuda oko nas i koriste se u razne svrhe, od stvaranja vizuelnih prikaza podataka da se impresioniraju potencijalni investitori, do izvještaja o napretku. Mnogi od alata za vizualizaciju podataka su otvorenog koda. Aplikacije se mogu koristiti u kombinaciji sa nekim drugim, ili sa već postojećim dizajn aplikacijama, koristeći JavaScript, JSON, SVG, Python, HTML5 ili drag-and-drop bez programiranja. Sledi prikaz najpoznatijih i najsveobuhvatnijih alata koji rade sa velikim podacima:

U interaktivnoj 3D vizualizaciji, korisnik može pratiti i analizirati na hiljade elemenata podataka. Program ITscape vizualizuje istorijsku seriju koju spaja sa aktuelnim podacima u realnom vremenu, gde se mogu videti na hiljade promenljivih istovremeno, mogu se interaktivno analizirati stavke ili grupe, i zumirati sve što deluje interesantno.



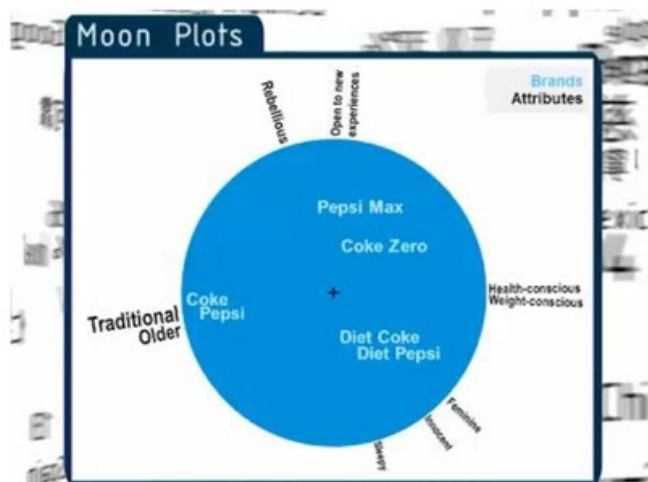
Slika 12. 3D Vizualizacija - ITscape [22]

Polymaps je besplatan Java Script program koji prikazuje kompleksne skupove podataka preko mapa.



Slika 13. Polymaps vizualizacija [23]

Q Research Software je moćna baza podataka i za istraživanje i vizualizaciju podataka, uglavnom služi kao sredstvo za pripremu izveštaja za istraživanje tržišta. Pogodan je za eksportovanje u Word, Excel, Power Point, CVS fajlove i PDF datoteke. Podržava višestruke tipove grafikona, histogram i tačkaste dijagrame. Takođe ažurira podatke u realnom vremenu i vrši statistička testiranja.



Slika 14. Q Research Software [24]

SAS Visual Analytics je alat za prikazivanje skupova podataka svih veličina i sveobuhvatne analitike. Sa intuitivnom platformom i alatom za automatsko predviđanje, SAS omogućava i početnim korisnicima da istražuju dublje odnose podataka i otkriju skrivene mogućnosti. Ključne karakteristike su te što može da se primeni u cloud sistemu, "Drag-and-drop autocharting" bira najbolji raspored za podatke, identifikuje potencijalno značajne korelacije, prilagođen je i kao mobilna aplikacija, kreira interaktivne izveštaje itd.



Slika 15. SAS Visual Analytics [25]

5.2. Uticaj Big data na edukaciju

Big data talas polako dostiže obrazovni sistem, posebno kroz preduzetničke inicijative koje nude širok spektar različitih učenja i sistemskih rešenja. Fakulteti danas imaju uvid u veliki broj podataka koji mogu da analiziraju i iskoriste. Iz aplikacije univerziteta i iz aplikacije za upis studenata, moguće je izvući podatke kao što su: lokacija, prethodne aktivnosti, zdravstveni problemi (fizički i emocionalni / mentalni), aktivnost u učenju, prethodne ocene, sadašnje ocene i status, socio - ekonomski podaci (prihodi), roditeljski status itd. Razvoj personalizovanih i adaptivnih sistema učenja je sadašnji hit obrazovnih skupova i upravo to može da bude ključ, ne da se student samo angažuje, već i da ima sistem koji može da odgovori na stvarne potrebe učenja svakog učenika. Štaviše, Big data softver provajderi grade sisteme koji nude informacije svim članovima obrazovne zajednice, koje mogu dovesti do efikasne saradnje. Kreatori politike predviđaju mogućnost zasnivanja svojih obrazovnih odluka na osnovu podataka u realnom vremenu prikupljenih iz onoliko mesta, koliko oni žele. Tradicionalne obrazovne institucije su snažni investitori u Big data. Glavni obrazovni izdavači okreću svoje napore ka dinamičnoj mrežnoj platformi koja je opremljena za prikupljanje podataka od studenata koji su u interakciji sa njima, pružajući prilagodljive i

prilagođene odgovore. *Pirson i McGraw-Hill* su nedavno udružili snage sa mladim kompanijama za adaptivno učenje kao što su *Knepton* [26] i *Aleks* [27]. Infrastrukturni softver kao što je *Blackboard* [28] i *Ellucian* [29] za visoko obrazovanje, baziraju svoje sisteme na alatima analitike podataka za predviđanje uspeha studenta na osnovu podataka prijavljenih od strane softverskih sistema njihovih klijenata, promovišući upotrebu Big data za pomoć pri poboljšanju ishoda učenja.

Preduzetničke inicijative povlače obrazovne industrije da budu inovativne i koristeći analize podataka izazivaju značajne pedagoške promene. Mnoge istraživačke grupe, pokreću projekte kroz istraživanje tehnika i modela iz veštačke inteligencije i nauke učenja. Multidisciplinarni projekat razvija tehnologije za analizu i podršku učenju preko okruženja poboljšanog različitim tehnologijama, kako unutar tako i van učionice. U kontekstu različitih vrsta sveprisutnih društvenih medija (Facebook, Twitter) koristi se skaliranje grupe za učenje od malih do veoma velikih grupa. Pokušavajući da integriše najnovije tehnike perceptivnog računarstva (vizuelno - praćenje oka ili 3D prepoznavanje gesta, govora, prepoznavanje emocija, itd.) u sisteme adaptivnog učenja, istraživački projekat u Intelu istražuje mogućnosti novih tehnologija na velikim podacima, prediktivnoj analitici i perceptivnom računarstvu kao izvoru informacija. Projekat namerava da stvori sistem koji kombinuje podatke koji se odnose na fizičku interakciju učenika sa računarom i kontekstualne podatke iz sistema podučavanja. Korišćenjem analitike učenja, sistem će obezbediti informacije koje odgovaraju svakom studentu. Personalizacija je zasnovana na skupu podataka i odgovora u realnom vremenu, dok je učenik u interakciji sa lekcijom.



Slika 16. Popularna ilustracija korišćenja Big data [30]

5.2.1. Analitika učenja

Obrazovanje koristi analitiku učenja, data mining ("rudarenje", "kopanje" podataka) i proces generisanja delotvornog znanja iz ogromnih količina podataka. Suštinska ideja iza analitike učenja je da se upotrebe podaci analize, da se prilagodi nastava individualnim potrebama učenika u realnom vremenu, na isti način na koji Amazon i Google koriste metriku da prilagode preporuke i reklame potrošačima. Analitika Big data je proces ispitivanja velike količine podataka u nastojanju da se otkriju skriveni obrasci ili nepoznate korelacije. Programeri aplikacija za analitiku obično koriste mali uzorak podataka u "pseudo-cloud" okruženju i primenjuju aplikaciju u cloud okruženju velikih razmera sa znatno više procesorske snage [31]. Analitika učenja omogućava predviđanje budućnosti učenika (na osnovu prethodnih obrazaca učenja), preporuku i pružanje povratne informacije prilagođene odgovorima učenika, personalizaciju opcija učenja, i prilagođavanje nastave i stilova učenja. Često se preklapajući sa konceptom analitičkog učenja, "kopanje" obrazovnih podataka je orijentisano na razvijanje načina za otkrivanje obrazaca u podacima kroz istraživanje, u potrazi za novim znanjem. Istraživči analitike učenja i kopanja podataka u potrazi su za aplikacijama koje koriste učenicima, kao i za unapređenjem informisanja. Bilo da su podaci o obrazovanju preuzeti iz učeničke upotrebe interaktivnih okruženja za učenje, kompjuterski podržanog kolaborativnog učenja, ili administrativnih podataka škola i univerziteta, oni imaju više nivoa smislene hijerarhije. Pitanja vremena i konteksta takođe igraju važnu ulogu u proučavanju obrazovnih podataka. Individualni podaci se analiziraju i stvaraju profile učenja, dijagnostifikuju snage, slabosti i izazove, kao i korišćenje odgovarajućih šablona za učenje [32].



Slika 17. Blackboard sveobuhvatni izveštaj [33]

5.2.2. Personalizovano i adaptivno učenje

Koncept personalizovanog učenja primorava tradicionalnu nastavu i provajdere sadržaja da traže načine da se fokusiraju na iskustvo studenta. Danas, sistemi koji rade u pravcu razvoja personalizovanog učenja imaju osnovu u ogromnim količinama podataka generisanih od strane učenika dok su u interakciji sa onlajn okruženjima za učenje, otkrivajući šta znaju i kako uče najbolje [34]. Sistemi su u stanju da analiziraju ove podatke i da preporučuju u realnom vremenu šta i kako student treba da prouči sledeće. Adaptivno učenje se široko koristi da označi adaptivne programe koji nude različite sadržaje učenicima, na osnovu procene onoga što izgleda da znaju. Njegova specifičnost u odnosu na opšti koncept personalizovanog učenja je da adaptivno učenje podrazumeva stalni proces prilagođavanja sistema baziran na stalnom dobijanju novih podataka. Adaptivne platforme za učenje kontinuirano prikupljaju podatke studenta u cilju da sistem uči i prilagodi put učenja studenta tako da se menja i poboljšava tokom vremena. Personalizovano učenje može takođe da uključuje sisteme izvedene iz metoda zasnovanog na pravilima stabla odlučivanja koji dovodi do unapred utvrđene staze. Adaptivno učenje pruža studentima modularno okruženje za učenje, što znači da je nastavni plan i program podeljen i individualizovan. Svaki student vidi drugačiji nastavni program adaptiran i prilagođen u skladu sa njegovim/njenim sposobnostima učenja i tempa, hvatanjem i beleženjem studentskih podataka svakim pritiskom tastera, i nudi ono

što bi učenik trebalo da radi sledeće. Podešavanjem sa svakom interakcijom radi održavanja odgovarajućeg izazova, učenici ostaju u njihovoj optimalnoj zoni učenja i omogućeno im je da razviju svoj puni potencijal za učenje. *Dreambox Learning* [35] uvodi okruženje za adaptivno učenje za matematiku za osnovne škole. Učenik uči u okruženju nalik avanturističkoj igri gde učenici pokazuju svoj rad i razmišljanje. Kada se govori o adaptivnom učenju, misli se na sistem koji se neprekidno prilagođava da reaguje u realnom vremenu na performanse svakog pojedinca i njegovu aktivnost u sistemu. Ono maksimizira verovatnoću da će student razumeti određeni koncept preporučujući pravu instrukciju, u pravo vreme.

5.2.3. Obrazovni podaci

Aktivnosti studenata pri učenju mogu se snimiti koristeći sisteme za upravljanje učenjem (Learning Management Systems - LMS). Podacima se upravlja na svim nivoima (pojedinaac, škola, država), u mnogim različitim sistemima u svim oblicima (strukturirani i nestrukturirani podaci) kroz tekstove dokumenata, slike, video snimke, interaktivne akcije, itd. Glavni izazov je da se integrišu različiti delovi podataka. Sve vrste studentskih podataka iz programa za testiranje i nastavnog softvera za ocenjivanje se čuvaju. Sistemi obrazovnih podataka uključuju sledeće životne aktivnosti učenika: odjavljivanje iz biblioteke, posete teretani, sportska učešća, posete kafeima i kupovine u knjižarama, trajanje studentskih sastanaka, vreme u i van studentskih domova, LMS prijave i sesije, blog i forum istorija komentara, korišćenje interneta, e-mailovi koji se šalju i primaju preko univerzitetskog e-mail naloga, stranice koje učenici čitaju u digitalnim udžbenicima, odlomci koje oni ističu, njihovi profili društvenih mreža, njihove posete vikipediji, video gledan na MOOC-u itd. [36].

5.3. Primeri primene Big data

5.3.1. Uticaj na nauku

Nauka doživljava veliki razvoj i istraživanje je trenutno uslovljeno novim mogućnostima koje Big data donosi. Makrosistemi, ili "velika ekologija" kako ih naučnici u NASA nazivaju, postaju mogući samo uz korišćenje ogromnog broja podataka. Ovi podaci mogu poslužiti za modelovanje složenih fenomena iz realnog sveta. CERN, švajcarska laboratorija nuklearne fizike, koristi procesorsku snagu distribuiranih računara u preko 150 centara širom sveta za analizu podataka za otključavanje tajni našeg univerzuma. Na Univerzitetu Kornel, Hod Lipson i Majkl Šmit, naučnici razvijaju složene programe veštačke inteligencije za primenu u robotici [37].

5.3.2. *Primena u sportu*

S obzirom na to da se u sportu danas ulažu ogromna finansijska sredstva, tehnologija Big data je i tu našla svoje mesto. Big data se primenjuje u fudbalu, košarci, bejzbolu, formuli, tenisu itd. Video analitika prati uspešnost svakog igrača koristeći senzorsku tehnologiju u sportskoj opremi, loptama ili golf štapovima (pružanje povratnih informacija preko pametnih telefona i cloud servera). Treneri koriste smart tehnologiju za praćenje ishrane i sna, kao i društvene medije za praćenje emocionalnog stanja sportista. U poslednje vreme u sportu se može uočiti široka primena senzora (zbog pada njihove cene), savremenih sistema, analitike, algoritama, mobilne tehnologije, alata za vizuelizaciju itd. [38].

U košarci se beleži konstantan napredak implementacije savremenih tehnologija. NBA ekipe uglavnom koriste SportVU [39] sistem koji je nastao u američkoj vojsci. Kamere snimaju utakmicu i u realnom vremenu trener na svom tabletu može da vidi statistiku koja može da mu bude od velike pomoći. Zatim tehnologija STATS predstavlja svetskog lidera u ovoj oblasti. Koristi se na primer u formuli 1. U fudbalu se na primer ugrađuju senzori u kopačke, odeću, na stadione itd. Ti čipovi prate 200 različitih podataka u realnom vremenu i sve podatke šalju treneru na tablet. Time se stiče poboljšanje performansi igrača i ekipe, ali se time mogu obogatiti i iskustva gledalaca. Senzori se masovno koriste i u tenisu. Na primer IBM je na Rolan Garosu omogućio da navijači mogu da imaju uvid u detaljnu statistiku sa turnira odigranih u proteklih 8 godina (po jednoj utakmici se beleži 41 milion podataka).



Slika 18. Aplikacija za prikaz statistike NBA igrača

(Izvor: www.stats.com)

5.3.3. *Primena u zdravstvu*

Budućnost zdravstva je da u bolnici osim doktora rade i informatičari, analitičari i statističari. Pomoću analitike se može unapred otkriti koji to pacijenti imaju povećan rizik od oboljevanja (dijabetes, srčane bolesti, astma, prevremeno rođena deca itd.). Nutricionisti primenom Big data rešenja dolaze do relevantnih zaključaka koje uspešno koriste u svojoj oblasti. Uveliko se grade baze DNK kako bi se pronašli obrasci ponavljanja i time predvidelo oboljenje pacijenata. Korišćenjem Big data može se uočiti mogućnost primene određenih lekova za lečenje nekih bolesti za koje se do tada nije znalo da ih ti lekovi mogu lečiti [40]. Analitika Big data omogućava praćenje i predviđanje kretanja epidemije i pojave bolesti. Na primer epidemija gripa se može detektovati u realnom vremenu integrisanjem podataka iz medicinske dokumentacije sa društvenih medija (od analitike šta ljudi kucaju, npr. "danas sam u krevetu – osećam groznicu").

Reg4ALL je inicijativa koja promoviše građanske akcije. To je platforma koja omogućava pojedincima i zajednicama da skupe podatke stvarajući otvorenu bazu podataka javnog zdravlja, koja omogućava lekarima da definišu efikasne tretmane odnosno lečenja.



Slika 19. Big data u medicini [41]

5.3.4. Još neki primeri upotrebe Big data

Google Search Engine je najbolji primer Big data koji svako može da koristi za pronalaženje određenih informacija. Prema nekim istraživanjima Google ima "najveću bazu podataka na planeti" [42]. Lični prenosivi gedžeti, kao što su pametni satovi ili smart narukvice, generišu podatke i obaveštavaju o tome kako naše telo funkcioniše na osnovu analize skupljenih podataka. Većina onlajn sajtova za upoznavanje primenjuje alate i algoritme Big data da pronađe najpogodnije partnere. Google Ngram Viewer omogućava praćenje kulturnih trendova kroz vreme koristeći pretragu određene reči. Alat se zasniva na ogromnom skupu podataka dobijenih iz miliona knjiga koje je Google digitalizovao.

Optimizacija saobraćajnih tokova je bazirana na informacijama o saobraćaju u realnom vremenu, podacima sa društvenih medija kao i vremenskim podacima. Trenutno postoje pilot projekti koji koriste analitiku Big data za razvoj inteligentnih gradova, gde su saobraćajne infrastrukture i komunalni procesi udruženi: autobus će čekati voz koji kasni, saobraćajni znaci će se menjati zavisno od realne situacije kako bi se smanjile gužve. Primer je implementacija Intelove Apache Hadoop tehnologije kao pomoć u prenaseljenim gradovima u Kini [43].

Američka vlada ulaže značajna sredstva u poboljšanje bezbednosti omogućavajući onima koji sprovode zakon, kao na primer NSA, da koriste analitiku Big data da osujete terorističke zavere ili da otkriju i spreče sajber napade.

Da bi se stvorili prognostički modeli potrebno je shvatiti ponašanje i preferencije kupaca. Zbog toga se prikupljaju podaci sa društvenih medija, logovanja, korišćenih poruka, senzora itd. Robne kuće su u stanju da veoma precizno pronalaze i definišu svoju ciljnu marketing grupu. Postoji čuvena ilustracija slučaja oca koji se naljutio na "Target", jer je njegova ćerka primila reklame u vezi trudnoće; on je tada saznao da je "Target" "znao" da je ona bila trudna pre njega, na osnovu podataka o njenoj nedavno promenjenoj navici pri kupovini kozmetike [44].

Optimizacija poslovnih procesa je moguća na osnovu predviđanja generisanih podataka kao što su podaci sa društvenih medija i veb trendovi sa pretraživača. Drugi primer je geografsko pozicioniranje i senzori identifikacije radio frekvencije koji se koriste za praćenje robe ili dostavnih vozila i optimizovanje maršrute integrisanjem podataka u realnom vremenu o saobraćaju.

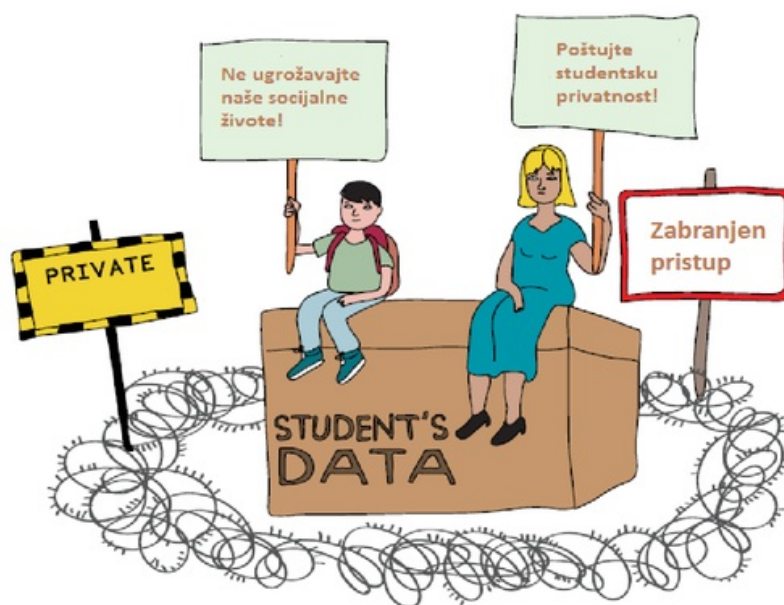
Profesor Sebastijan Trun (*Stenford Univerzitet*) i istraživač Piter Norvig (Google) vode projekat za izgradnju autopilot vozila koje koristi algoritme veštačke inteligencije i podatke prikupljene iz snimaka i merenja Google Street View vozila.

5.4. Skepticizam

Skepticizam je takođe važan deo priče oko Big data. Stručnjaci se slažu da još uvek postoji dug put do momenta kada se složeni podaci koji se generišu mogu lako transformisati u smislenu informaciju. Neki od problema su:

- Podaci su često nesređeni, nestrukturirani i oštećeni.
- Teško je odabrati relevantne podatke.
- Nedostatak poslovne podrške.
- Troškovi infrastruktura.
- Nedostatak dovoljnog znanja i profesionalaca koji se ovim bave i tumače složene i raznovrsne multi-sisteme [45].

Kritičari ukazuju da je učenje vođeno podacima (netradicionalno učenje) pretnja koja može da pretvori škole u fabrike. Razna istraživanja širom sveta o onlajn privatnosti i svesti o korišćenju podataka pokazuju zabrinutost stanovništva zbog ugrožene onlajn bezbednosti.



Slika 20. Zanimljiva ilustracija povodom ugrožavanja privatnosti studenata [45]

Glavni internet provajderi su već otvoreno raspravljali o "trgovanju" podacima korisničkih profila čiji su oni vlasnici i kojima upravljaju. Od Google-ovih 37,9 milijardi dolara prihoda u 2011. godini, 96% je zarađeno od oglašavanja [46]. Postavlja se pitanje da li Google pruža školama slobodan pristup preko aplikacije za obrazovanje Google Apps u nadi da će na kraju zaraditi od oglašavanja i kopanja podataka uskladištenih školskih zadataka studenata i interakciji u vezi sa obrazovanjem. Na početku 2013. godine, Masačusets je postao prva američka država koja brani kompanijama koje pružaju usluge cloud computinga da studentske podatke koriste u komercijalne svrhe. Nefitne finansirajuće platforme poput *inBloom* očekuju da počne naplaćivanje okruzima pokretanje korišćenja infrastrukture 2015. godini (2-5 dolara po učeniku godišnje). Pored toga, provajderi aplikacija će tražiti preraspodelu dobitaka. Sve ovo dovodi do sledećih pitanja: "Ko poseduje iskustvo učenja? Ko poseduje sve podatke o obrazovanju? Kompanije? Škole? Instruktori? Studenti? Da li učenici znaju da se o njima prikupljaju i skladište podaci? Kako možemo biti sigurni da analitika učenja i "kopanje" podataka ne umanjuje vrednost? Kako biti siguran da učenik nije samo predmet analize? Kako proveriti da li student ima kontrolu nad svojim podacima i njihovom izučavanju?"



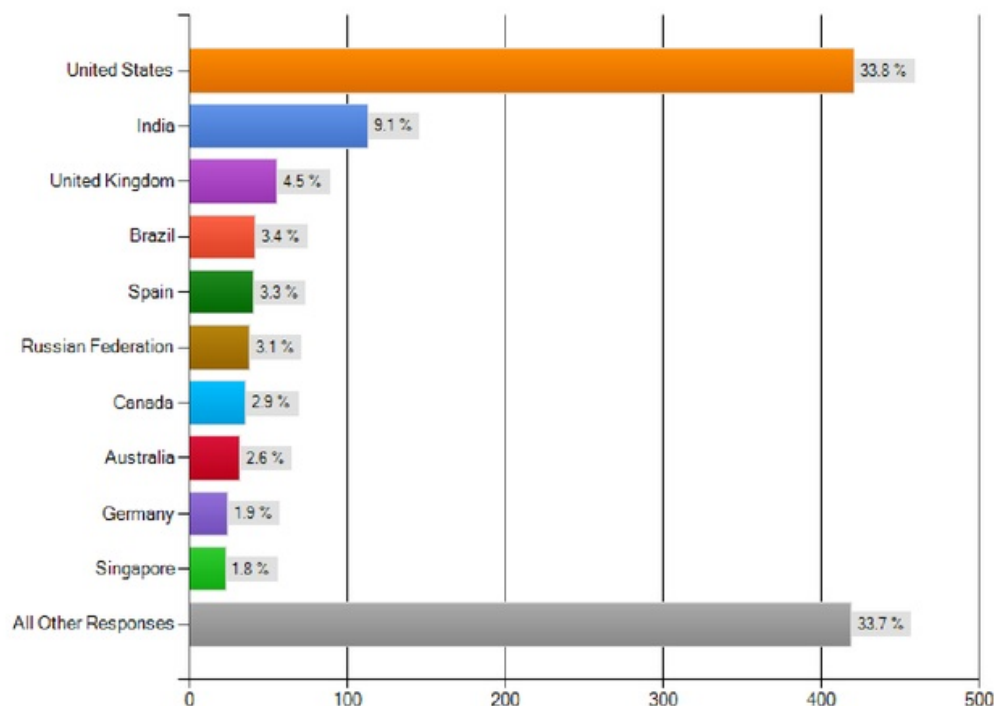
Slika 21. Ilustracija vrednosti informacija [46]

6. MASOVNI OTVORENI ONLAJN KURSEVI (MASSIVE OPEN ONLINE COURSES - MOOC)

U ovom delu rada dat je prikaz najpoznatijih platformi za učenje na daljinu koji se koriste u otvorenim masovnim onlajn kursovima. MOOC predstavlja projekat na kom rade najpoznatiji svetski univerziteti i on je u stalnom razvoju. Mooc izgrađuje zajednice studenata i profesora kao i učenika i nastavnika, obezbeđuje mnoštvo foruma gde se diskutuje na određenu temu. Kroz razne materijale za čitanje i video klipove MOOC doprinosi usavremenjivanju tradicionalnih kurseva.

MOOC je postao popularan 2008. godine i to uglavnom u Americi. 2011. godine Univerzitet Stendford je započeo tri probna onlajn kursa koji su imali po 100000 prijavljenih korisnika. Nakon toga, profesori sa tog univerziteta su osmislili internet sajt koji je specijalizovan za tu vrstu kursa i nazvali su ga Coursera. 2012. godine je dosta univerziteta uvidelo kakve sve mogućnosti može da pruži ovakav vid učenja, i Harvard univerzitet formira svoju platformu za onlajn kurseve koja se zove edX.

U grafikonu 2. je prikazana statistika iz kojih zemalja su uglavnom polaznici MOOC kurseva. Nema preciznih podataka koliko je polaznika bilo iz naše zemlje, ali sudeći po porukama na forumima sa platforme Coursera, moglo bi se reći da je broj od nekoliko stotina.



Grafikon 2. Zemlje iz kojih su polaznici MOOC kurseva [47]

6.1. Pregled poznatih MOOC platformi

Tri najpoznatije MOOC platforme su Coursera, Udacity i EdX. Coursera [48] je najinteresantnija upravo zbog toga što nudi kurseve iz puno različitih oblasti. U početku su svi kursevi bili besplatni a sad je već uvedeno plaćanje za naprednije kurseve ili za sertifikat koji može da se dobije nakon položenog testa. Udacity platforma je uglavnom bazirana na kurseve iz matematike. Kursevi su raspoređeni od najlakših ka najtežim i u zavisnosti od prethodnog znanja, polaznik kursa bira koji će kurs odslušati. EdX platforma je najpoznatija kad se priča o oblasti informatike ali takođe nudi kurseve i iz drugih oblasti. Zajedničko za sve je da polaznik dobije odgovarajuću literaturu i izabere vreme kad će odslušati određeni kurs. Vrlo često ove MOOC platforme nude naprednije kurseve koji nisu besplatni, a takođe nude i sertifikate koji su naširoko priznati od velikih svetskih kompanija.

U tabeli 2. su prikazane neke od prednosti i mana 3 najpoznatije MOOC platforme [49].

	Prednosti:	Mane:
COURSERA	- Ogroman broj kurseva iz najrazličitijih tematskih oblasti - Puno informacija o kursu - Video uvod je napravljen za sve kurseve - Veliki broj kurseva ima mogućnost učitavanja prevoda na različitim jezicima - Uglavnom se nudi sertifikat za položen kurs	- Nisu dostupni svi kursevi - Postoji ocenjivanje nekih vežbi koje se mogu odbiti - Moraju se ispuniti svi rokovi kao uslov za sertifikat - Procena nedeljnog fonda časova je ponekad nerealna
UDACITY	- Mogućnost prijave sa Facebook ili Twitter nalogom - Svaki kurs ima video pregled - Moderni kursevi - Interaktivni tutoriali - Korisnik sam određuje tempo učenja, ne postoje rokovi - Svi kursevi nude sertifikate - Odličan forum	- Veliki kursevi - Nedovoljno popunjen opis kurseva - Ne postoji prevod
EDX	- Dobar izbor kvalitetnih kurseva - Puno informacija o kursu - Uglavnom kursevi imaju video uvod - Različiti jezici - Sertifikati - Forumi izabranog kursa	- Nisu dostupni svi kursevi - Neki kursevi za koje je potrebno predznanje nude ocenjivanje polaznika - Procena nedeljnog fonda časova je ponekad nerealna

Tabela 2. Poređenje platformi Coursera, Udacity i Edx

	Maksimalna veličina razreda	Brendiranje	Analitika polaznika	Monetizacija	Mobilni	Hosting
	300.000	✓	✓	✓	✓	Samostalni Hosting
	10.000	✓	✓	✓	✓	Samostalni Hosting preko trećeg lica
	Neograničeno	✗	✓	✗	✓	Hostovanje
	Neograničeno	✗	✗	✓	✓	Hostovanje
	Neograničeno	✗	✗	✗	✓	Hostovanje

Slika 22. Uopšteno i ukratko poređenje još nekoliko MOOC platformi.

6.2. Povezanost MOOC-a i Big data

Prikupljanje masovnih obrazovnih podataka na tržištu je poraslo eksponencijalno u poslednjih nekoliko godina. MOOC je usmeren na neograničeno učešće i otvoreni pristup preko veba, koristi kratka video predavanja u kombinaciji sa nizom procena prilagođenja masi studenata (preko 100 hiljada studenata na popularnim kursevima), automatizovane povratne informacije putem onlajn procene (na primer, kvizovi i ispiti) i recenzije i aktivnosti grupne saradnje. Coursera, na primer, prikuplja podatke za svaku akciju (ili neaktivnost) koju obavlja student - kada student pauzira video, povećava brzinu reprodukcije, odgovara na pitanje testa, revidira zadatak ili komentare u forumu. Ovaj mikroskopski nivo podataka, kada se podigne na skalu na kojoj MOOC radi, olakšava identifikaciju podrazumevanu u sistemu. Kako Dafne Koler, suosnivač Coursera, ističe: "Ako dva učenika u klasi univerziteta od sto daju pogrešan odgovor, vi nikada ne bi primetili, ali kada dve hiljade ljudi da isti pogrešan odgovor, to je nekako teško ne primetiti" [50]. Ogromne količine generisanih podataka od upisanih kurseva, otprilike između 10.000 i 100.000 polaznika, omogućavaju provajderima da poboljšaju rezultate kao što su optimizacija materijala. Ako je test pitanje odgovoreno netačno, ili ako studenti izgube fokus tokom određene tačke u toku kursa, podaci mogu pomoći kreatoru da se vrati i u nastavni program nešto doda ili izmeni. MOOC provajderi iskorišćavanjem ove skale eksperimentišu sa nastavnim materijalom za kurs, metodama prezentacije i komunikacije sa studentima.

7. UTICAJ I PRIMENA DRUŠTVENIH MREŽA U NASTAVI

Kao rezultat uticaja informacionih tehnologija na društvo, dolazi do toga da se sve više koriste i primenjuju društvene mreže kako u poslovanju tako i u školovanju. Sve je više učenika, studenata, profesora, nastavnika i roditelja koji su uključeni u društvene mreže kao što su Twitter, Facebook, Youtube, LinkedIn itd. Primena u obrazovanju i školskoj komunikaciji dovodi do toga da se granice između obrazovnog i ličnog umrežavanja polako gube. Da su društvene mreže postale trend u nastavi potvrđuju i sledeći podaci:

- 96% studenata s pristupom internetu izgrađuje društvenu mrežu
- 60% studenata na društvenim mrežama razgovara o temama vezanim za obrazovanje
- 50% tinejdžera tvrdi da onlajn (preko Facebook-a, blogova itd.) ili putem sms poruka, razgovaraju sa vršnjacima o školi [51].

Prednosti koje pružaju društvene mreže su drugačije od fakulteta do fakulteta jer različiti fakulteti koriste društvene mreže drugačije. Neki fakulteti ili škole koriste društvene mreže isključivo za komunikaciju sa studentima, a neki ih koriste kako bi unapredili učenje s obzirom na činjenicu da društvene mreže doprinose studentima lakše povezivanje u obliku virtualne zajednice. Neka istraživanja pokazuju da će u roku od najmanje deset godina 70% budućih generacija nastavu pohađati u virtualnim školama gde će se i vežbe i ispiti održavati na vebu. U tom slučaju će profesori iskoristiti društvene mreže Facebook, Twitter ili Youtube da angažuju studente ili učenike u nastavi. Pošto je ustanovljeno da se društvene mreže mogu koristiti u edukaciji, uvedene su neke pozitivne promene na području obrazovanja koje bi studentima donele više motivacije, saradnje i kreativnosti.

Kako je internet postao svakodnevica i kod profesora i kod studenata, skočili su i standardi u korišćenju informacionih tehnologija u obrazovanju. Profesoru nije više dovoljno da ima u kabinetu računar i da jednom dnevno proveri elektronsku poštu, već bi trebalo da u svakom trenutku prati razvoj studenata tako što bi deo nastavnog sadržaja prebacio u stabilan internet prostor gde bi se kreirao zajednički prostor sa studentima. Tu bi se nalazio i elektronski dnevnik gde bi studenti, kao i roditelji, u svakom trenutku mogli da imaju uvid u nastavne aktivnosti i ocene. Predavač bi svoj

sistem edukacije upotpunio a menadžment škole bi imalo direktnu komunikaciju sa zaposlenima u smislu razmene dokumenata, aktuelnih i dopunjenih rasporeda, najnovijih obaveštenja itd.

Društveni softver (eng. social software) omogućava kreiranje i održavanje postojećih društvenih veza između ljudi, korišćenjem računarskih sistema. Istraživanja u ovoj oblasti obuhvataju različite discipline uključujući matematičke alate i tehnike, teoriju igara i računarstvo. Društveni mediji (eng. social media) predstavljaju ljudske interakcije korišćenjem virtuelnih zajednica i mreža, kreiranje, diskusiju i modifikaciju korisnički generisanog sadržaja.



Slika 23. Društvene mreže

(Izvor: <http://blog.erazvoj.com/?p=12569>)

Društvene mreže su postale svakodnevnicu u svim sferama života pa i u obrazovanju. Podaci koje je objavio *Global Web Index* pokazuju da Facebook trenutno ima više od milijardu aktivnih korisnika, dok se na drugom mestu nalazi Google+ koji ima oko 400 miliona korisnika. Objavljen je i podatak da sve vodeće društvene mreže beleže porast korisnika, Facebook za 33%, Google+ za 27% i Twitter za 40%.

Osnovna karakteristika društvenih mreža je kolektivna inteligencija, fenomen koji se naziva i „mudrost gomile“. Čim veliki broj ljudi iz različitih delova sveta, različite vere, godina i

obrazovanja, priča o nekoj temi, proces i rezultati analize pokazuju bolje i relevantnije rezultate od procesa kada se gledaju odvojeno mišljenja pojedinca [52].

Dobar primer ovog fenomena je i takozvano tagovanje. Ovo klasifikovanje postiže se kreiranjem i upravljanjem tagovima. Kreiranje tagova ili „tagovanje“, predstavlja vrlo popularan način klasifikovanja sadržaja. Tag predstavlja ključnu reč čija je uloga označavanje predmeta ili kategorije. Korisniku je dozvoljeno da pridruži bilo koju ključnu reč kako bi identifikovao sadržaj, kao što je link u slučaju social bookmarking-a ili pesme, video ili fotografije u slučaju sajtova za razmenu digitalnog sadržaja. Bitno obeležje je da se tagovi mogu razmenjivati, koristiti u pretragama ili preporučiti na osnovu izbora drugih korisnika istog sadržaja.

Pojava sajtova društvenih mreža omogućilo je da veb postane najveći javni izvor podataka na svetu. Ti javni podaci su dostupni za analizu ljudima i njihovim međusobnim vezama. Oni omogućavaju identifikovanje najvažnijih ili najuticajnijih ljudi u mreži, kako se stvari šire kroz mrežu i prirodu ljudskih veza.

Radi boljeg pregleda korišćenja društvenih mreža u Srbiji, urađena je anketa nad 876 korisnika društvenih mreža. Anketu je sproveo sajt <http://www.adriatalk.com>.

5

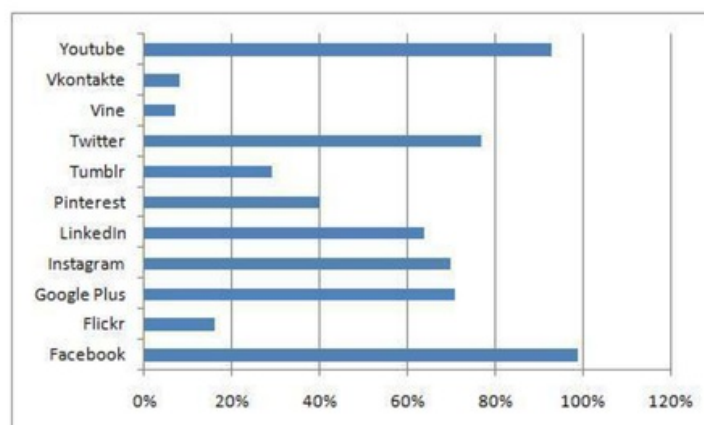
Pitanje 1 – Na kojim društvenim mrežama imate nalog?

<i>Facebook</i>	<i>864</i>
<i>Flickr</i>	<i>144</i>
<i>Google Plus</i>	<i>624</i>
<i>Instagram</i>	<i>612</i>
<i>LinkedIn</i>	<i>564</i>
<i>Pinterest</i>	<i>348</i>
<i>Tumblr</i>	<i>252</i>

<i>Twitter</i>	672
<i>Vine</i>	60
<i>Vkontakte</i>	72
<i>Youtube</i>	816

Tabela 3. Odgovori anketiranih korisnika na pitanje br 1.

(Izvor: <http://www.adriatalk.com>)



Grafikon 3. Odgovori na pitanje 1.

5

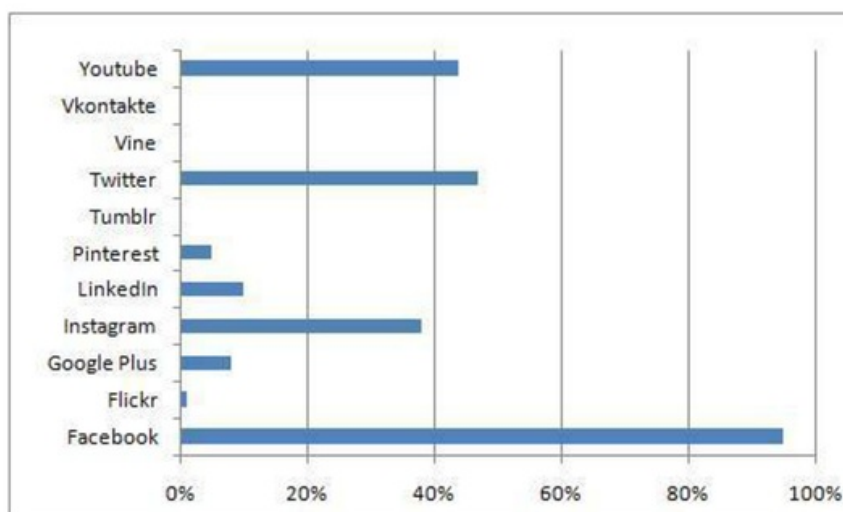
Pitanje 2. Na kojim društvenim mrežama provodite najviše vremena?

<i>Facebook</i>	828
<i>Flickr</i>	12
<i>Google Plus</i>	72
<i>Instagram</i>	336

<i>LinkedIn</i>	84
<i>Pinterest</i>	48
<i>Tumblr</i>	0
<i>Twitter</i>	408
<i>Vine</i>	0
<i>Vkontakte</i>	0
<i>Youtube</i>	384

Tabela 4. Odgovori anketiranih korisnika na pitanje br 2.

(Izvor: <http://www.adriatalk.com>)



Grafikon 4. Odgovori na pitanje 2.

Kod drugog pitanja ispitanici su za odgovor mogli da odaberu najviše 3 odgovora. Google plus je dobio dosta pozitivnih odgovora na prvom pitanju zato što korisnici android telefona uglavnom imaju "po default-u" i profil na Google+ mreži. O tome govori i podatak da samo 8% ispitanika zaista koristi tu mrežu. Slične ankete su rađene i 2009. godine, i tada su rezultati bili isti kod jedne stvari - u Srbiji se ubedljivo najviše koristi Facebook društvena mreža. Kod nas je *IT Academy* [53] među prvima počela sa aktivnom upotrebom društvenih mreža u nastavi. Kreirani su razni oblici učenja na daljinu i uspešno su povezani sa komunikacijama na društvenim mrežama.

7.1. Facebook

Facebook je počeo sa radom 2004 godine. Prvobitno, članovi ove internet stranice bili su samo studenti sa Harvarda, da bi se kasnije to proširilo na druge studente u Americi. Danas ta društvena mreža predstavlja najpopularniju mrežu na svetu i broji oko 750 miliona registrovanih korisnika širom planete. Kao najpopularnija društvena mreža, Facebook se primenjuje u nastavi i informativno ali i edukativno u užem smislu te reči. Fakulteti najčešće na Facebook-u prave sopstvene stranice pomoću kojih obaveštavaju svoje pratioce, najčešće studente o servisnim i obrazovnim novostima, novonastalim situacijama i aktuelnostima u okviru nastavnog programa.

Primena Facebook-a u edukaciji može biti različita. Najčešće podrazumeva se oformljavanje grupe studenata određenog smera ili iz određene oblasti. Tada se vrši komunikacija ne samo predavača prema studentima jednosmerno, već i samih studenata međusobno. Što se tiče privatnosti grupe, administrator uglavnom bira da grupa bude zatvorenog tipa.

U poslednje vreme koristi se model po kom student ili učenik treba da se sam angažuje i u okviru nekog zadatka kreira neki oblik ili specifičnu formu nastavnog zadatka baš u okviru Facebook platforme. Na taj način nastaju fan stranice ili novi profili poznatih ličnosti (npr književnika) koje su kreirali studenti u okviru svojih studijskih obaveza. Slični pilot projekti se realizuju u beogradskim gimnazijama i to iz srpskog jezika, gde učenici imaju zadatke da naprave i održavaju profile naših poznatih književnika.

Može se reći da je Facebook kao društvena mreža postala jedna od platformi za učenje na daljinu. Profesori na Facebooku imaju mogućnost da objavljuju svoja predavanja u obliku Power Point prezentacije, ili mogu sastavljati razne vrste kvizova, igara ili nadopunjujućih multimedijских sadržaja. Student automatski aktivno učestvuje u kreiranju i korišćenju gradiva na interesantan i zanimljiv način.

7.2. Twitter

Ako se posmatra ranije navedena statistika o korišćenju društvenih mreža u Srbiji [54], može se lako zaključiti da je Twitter postao jedna od najpopularnijih socijalnih mreža u našoj zemlji. Svake godine se popularnost Twittera ²³brzo povećava i to dovodi do toga da ova društvena mreža može da pruži mnogo toga svojim korisnicima a naročito onima koji rade u obrazovanju. Ova mikroblog platforma je vrlo jednostavno smišljena. Korisnici objavljuju svoje tvitove tako što kucaju maksimalno 140 karaktera i svi drugi korisnici koji ih prate mogu to da vide. Mogućnost primene u edukaciji je velika i može se koristiti u najprostijem obliku npr. profesor postavlja pitanja, studenti odgovaraju i samim tim se ostvaruje komunikacija u realnom vremenu, a trag ostaje sačuvan. Mogu se kreirati razni online kvizovi. Ne treba zanemariti ni činjenicu da od svih društvenih mreža baš na ovaj način korisnici dobijaju najnovije i najpouzdanije informacije svih vrsta, direktno od izvora ili aktera svih sfera, a spremnih za dalju eksploataciju i transfer.

Twitter je pun raznovrsnih informacija. Sve poznate kompanije imaju twitter naloge kao i poznate ličnosti. Studenti iz bilo kojih oblasti mogu vrlo lako da dopunjuju i proširuju svoje znanje putem Twittera tako što imaju mogućnost da prate kompanije ili ljude čije mišljenje vrednuju i samim tim blagovremeno dobijaju informacije na njihovim listama. Twitter nema ograničenja kako se koristi, tako da su mogućnosti prikupljanja informacija ogromne. U drugom delu rada ova društvena mreža će se detaljnije opisati i biće prikazane ideje i realizacije koje su potpomogle osavremenjivanju nastavnog sadržaja u visokoškolskim ustanovama.

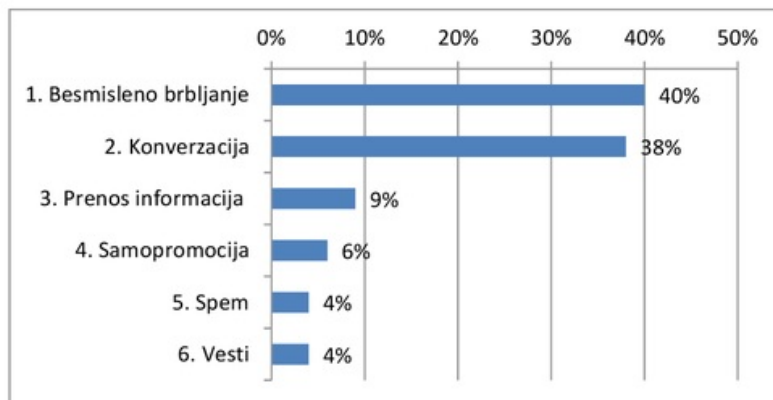
7.2.1. Istorijat Twittera i opis podataka

Twitter predstavlja društvenu mrežu koja omogućava korisnicima da šalju i čitaju kratke poruke od maksimalno 140 karaktera. Nastao je u martu 2006, a zvanično je pokrenut u julu 2006. Twitter se razvija neverovatnom brzinom, dostigavši preko 200 miliona korisnika sa preko 200 miliona tvitova dnevno. Korisnici kreiraju svoj nalog na Twitteru, i od trenutka kada dobiju nalog mogu početi da "tvituju", što je izraz za slanje poruka na Twitteru. Korisnici mogu da se prijave na tvitove drugih korisnika, što je proces poznat kao "praćenje". Ovi korisnici se zovu "pratioci". Tvitovi koje korisnik šalje su vidljivi svima, međutim korisnici mogu da odaberu opciju slanja tvitova posebno svojim "pratiocima" koji neće biti vidljivi za javnost.

Korisnici Twittera se identifikuju korisničkim imenom, a to korisničko ime praćeno je simbolom "@". Kada korisnik označi drugog korisnika u svom titu njegovim korisničkim imenom, to će biti vidljivo za javnost, a korisnik koji je označen će biti obavešten putem Twittera da je "pomenut". Ukoliko korisnik vidi tit koji mu se učini interesantnim i želi dalje da prenese tu informaciju, može

da "retvituje" taj post, što je slično prosleđivanju e-pošte novoj grupi korisnika, u ovom slučaju njihovih pratioca. Retvitovi u poruci će uvek biti označeni sa "RT". Na kraju, poruke se mogu grupisati po temama ili tipu upotrebom takozvanih heštagova (#). Heštag koji prethodi temi omogućuje Twitter korisnicima da pronađu tvitove koji se odnose na određenu temu prilikom vršenja pretrage [55].

Twitter takođe ima opciju prikazivanja lokacije. Ako korisnici tviuju sa mobilnog uređaja, oni mogu da izaberu da uključe opciju prikazivanja lokacije, a njihova geografska dužina i širina biće zabeležena u tviu. Tvitovi mogu da se odnose na bilo šta, ali veći deo sadržaja na Twitteru odnosi se na nekoliko ključnih kategorija. Te kategorije su navedene u istraživanju sprovedenom 2009. godine od strane Pear Analytics kompanije na osnovu uzorka od 2.000 tvitova [56]. Ova studija je pokazala da se tvitovi prvenstveno odnose na ovih šest kategorija (Grafikon 5):



Grafikon 5. Prikaz korišćenih kategorija

Iako se ove brojke odnose na istraživanje sprovedeno 2009. godine kada Twitter nije bio toliko rasprostranjen kao danas, ova opšta kategorizacija tvitova je i dalje na snazi. Ako se na primer govori o edukaciji, sfere interesovanja bi predstavljale kategorije 2, 3 i 6 što bi činilo 51% tvitova. Treba imati u vidu da neće svi tvitovi iz ovih kategorija biti od koristi u edukaciji. Bitno pitanje za istraživanje je na koji način analizirati tvitove i kako važne informacije odvojiti od nevažnih.

7.3. Upoređivanje Facebooka i Twittera

Twitter i Facebook se veoma razlikuju. Twitter je društvena mreža koja je zamišljena kao ogromna baza informacija i kreiran je sa ciljem da informacije brzo kruže. Da bi pročitali nečiju objavu, korisnici na Twiteru mogu da "prati" korisnika koji tviuje a i ne moraju. Sve je otvoreno i

zasnovano na tome da korisnik sam odlučuje kakve će informacije dobijati. Facebook je drugačija vrsta društvene mreže, bazira se uglavnom na komunikaciji i prijateljstvu između ljudi koji se već poznaju u životu. Namenjen je porodicama i ljudima koji se nisu dugo videli i pruža mogućnost ponovnog sklapanja prijateljstva. Uglavnom se deli puno fotografija a pisanje nije ograničeno kao kod Twittera. Facebook je nastao 2003. godine a Twitter 2006. godine. Zajedničko je to da su korisnici najčešće između 18 i 34 godine. U poslednje vreme Facebook beleži najveći rast korišćenja starijih od 55 godina a Twitter dvadesetogodišnjaka.

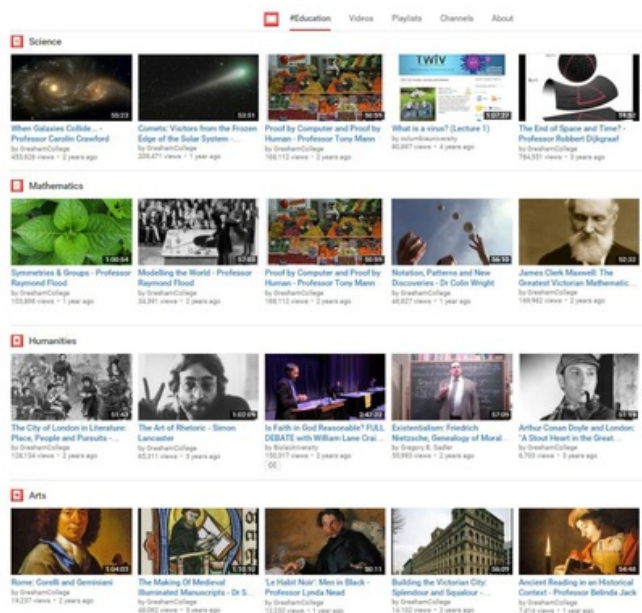
U anketi sprovedenoj preko interneta, učestvovalo je 300 ispitanika, od čega 70 odsto iz Evrope, a ostatak sa područja Severne Amerike i azijskih zemalja. Svi ispitanici odgovarali su na set pitanja o tome kako koriste Facebook i Twitter i koju od dve društvene mreže više koriste [57]. Da bi rezultate upitnika mogli da povežu s karakterom ispitanika, svaki ispitanik je radio test ličnosti – standardni merni instrument koji ličnost procenjuje na bazi pet bazičnih osobina: otvorenost (za nova iskustva), savesnost, ekstrovertnost, emocionalna (ne)stabilnost (neuroticizam) i sklonost ka saradnji. Dodatno, kod ispitanika je merena intelektualna radoznalost, potreba za kognitivnom stimulacijom ili sticanjem novih znanja. Otkriveno je da je društvena upotreba Fejsbuka (korišćenje ove mreže kao kanala za druženje i proširivanje kruga prijatelja) u korelaciji sa visokim skorom na testu socijalizacije i emocionalne stabilnosti kao izraženim crtama ličnosti. Ovo se uklapa u rezultate ranijih istraživanja, koja sugerišu da stidljivi ljudi često koriste internet i društvene mreže kako bi uspostavili kontakte, stekli poznanstva i pobedili usamljenost. Društvena upotreba Twittera povezana je s visokim skorom na testu socijabilnosti i otvorenošću kao dominantnom crtom ličnosti (ali ne i sa neuroticizmom), dok je u negativnoj korelaciji sa savesnošću. Ovakav rezultat sugerise da korisnicima Twitter ne služi za borbu sa usamljenošću, koliko za neku vrstu izvora informacija. Ljudi koji koriste Fejsbuk za razmenu informacija - traže i dele informacije na manje zahtevan način, izbegavajući članke u časopisima i obimnije novinske izveštaje. Ograničeni kratkim statusima, korisnici Twittera koriste ovu društvenu mrežu za kognitivnu stimulaciju, bez socijalizacije. Pojednostavljeno posmatrano, Facebook je više za druženje, a Twitter za deljenje i razmenu informacija. Pošto se u ovom radu istražuje način primene i uticaj analize društvenih mreža na edukaciju, izabran je Twitter kao glavni izvor informacija koje su aktuelne i koje mogu da poboljšaju kvalitet obrazovanja.

Kako bi se obogatilo sadržaj nastave (na predmetu Informacione tehnologije u biznisu), pored standardnog nastavnog plana, tema koja se obrađuje je i analiza društvenih mreža, gde studenti uče o značaju analize društvenih mreža i kako da izvode analizu Twitter društvene mreže korišćenjem tehnologije IBM BlgInsights. Pojavom društvenih medija analiza društvenih mreža je definisana kao nauka, upravo zato što se korišćenjem ovih sajtova svakodnevno na vebu pojavljuju velike

količine podataka, i što su ti podaci javno dostupni preko API-a (eng. application programming interface). Zbog toga se korišćenjem tehnika analize društvenih mreža dolazi do interesantnih novih informacija, koje mogu biti od značaja za različite oblasti, kao što su informatika, ekonomija, marketing, politika, sociologija itd. S obzirom da u Beogradskoj poslovnoj školi postoji 7 smerova koji uključuju ove oblasti, moguće je ideju sprovesti i dalje na druge odseke, gde bi i profesori koji nemaju informatička iskustva mogli da osavremene svoja predavanja najnovijim trendovima putem Twitter analize.

7.4. Youtube

Youtube predstavlja najveću društvenu veb stranicu u svetu koja je namenjena podeli video materijala, samim tim može imati mnogo uticaja i primene u edukaciji. Tipična upotreba videa je omogućila predavačima da zabeleže tj snime svoja predavanja, a zatim ih postavljaju na Youtube kako bi studenti mogli da ih pregledaju i to koliko god puta im je potrebno. Jedan od primera je Youtube EDU na kojem je moguće naći materijale i predavanja iz raznih oblasti, kao što su informatika, matematika, hemija, fizika, književnost, umetnost i slično. U Srbiji postaje sve učestalije postavljanje nastavnih materijala od strane nastavnika, kao i postavljanje snimljenih predavanja na YouTube.



Slika 24. Youtube u edukaciji

(Izvor: <https://goo.gl/T1fX81>)

Youtube se može svrstati u kolaborativne i interaktivne alate koji omogućuju predavačima i studentima razmenu mišljenja i razvoj diskusije na određenu temu. Metod koji se sve češće primenjuje u edukaciji i na fakultetima u poslednje vreme je video striming. Vrlo je popularan metod zato što omogućuje da svako može besplatno da postavi prenos nekog događaja (Na primer IT događaja, skupova, seminara...) a pritom sve to je moguće pratiti sa svih krajeva sveta u realnom vremenu pomoću Youtube video strminga.

7.5. Google Plus

Ovu društvena mreža je razvijena od strane kompanije Google 2011. godine. Pošto je Google sve svoje servise povezao, korisniku je pomoću Google Plus društvene mreže omogućeno skoro sve kao i na drugim najpopularnijim mrežama. Analizirajući rezultate raznih istraživanja na temu korišćenja društvenih mreža u Srbiji, može se zaključiti da ova mreža nije toliko zaživela u Srbiji. Veliki broj ljudi ima profil na ovoj mreži ukoliko ima otvoren Google korisnički nalog, ali Facebook je jednostavno nezamenljiv barem kod nas. Google Plus ima odličnu opciju da se pomoću Hangouts Talk-a može uključiti 10 osoba u video konferenciju, tako da je to jedan od mogućnosti da se koristi u edukativne svrhe. Postoje i video chat-ovi, i deljenje većih fajlova pomoću Google Drive.

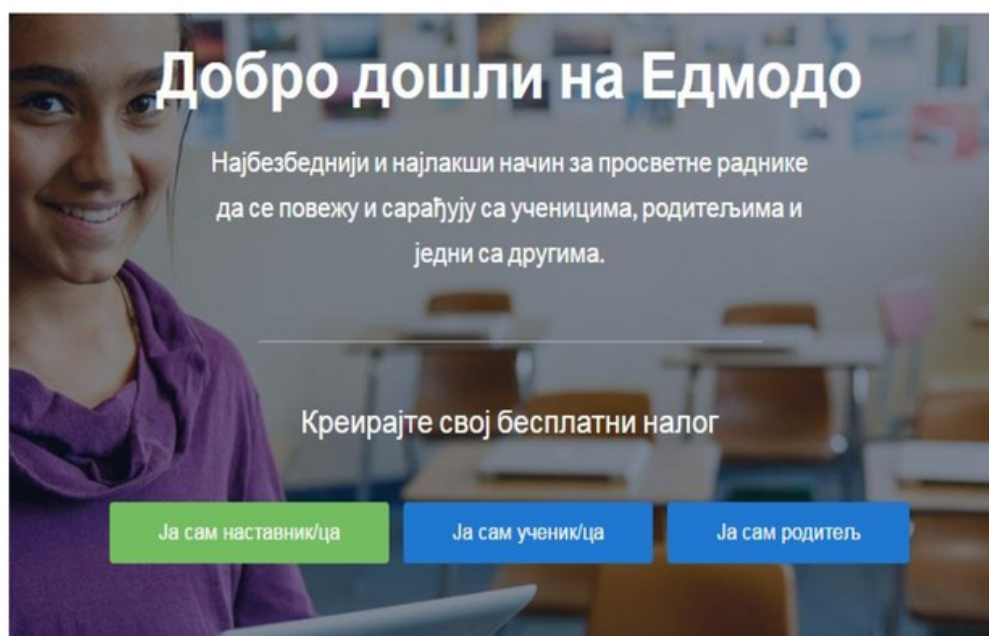
7.6. Edmodo

Potencijal uticaja društvenih mreža uvidele su i neke od društvenih mreža koje su isključivo namenjene za tu svrhu elektronskog učenja. Jedna od njih je i Edmodo, u Srbiji manje poznata društvena mreža ali sa velikim potencijalom da se primenjuje u edukaciji. Edmodo je obrazovna tehnološka kompanija koja nudi komunikaciju, saradnju i trening alate za škole i predavače. Edmodo mreža omogućava predavačima da dele sadržaj, distribuišu kvizove, zadatke, i upravljaju komunikacijama sa studentima, kolegama i roditeljima.

Mreža Edmodo je nastala 2008. godine. Godine 2014. Edmodo pokreće „Snapshot“ - komplet alata za procenu i merenje napretka učenika o obrazovnim standardima. Kompaniji se pridružuju i dva glavna izdavača u Britaniji „Okford University Press i Cambridge University Press“. Edmodo povezuje 51 milion učenika i nastavnika.

Edmodo istovremeno teži ka tome da se osavremeni nastavni program i izvrši povezivanje predavača i učenika i van učionice i lokacijskih okvira. Veoma podseća na Facebook, što dovodi do toga da je sama upotreba lakša, jer su uglavnom svi upoznati sa načinom funkcionisanja Facebook-a. Ova socijalna mreža nudi pravljenje testova i raznih zadataka, specijalizovanih kalendara, anketa

i blogova koji su od posebne važnosti za sam process edukacije. Ona pomaže u rešavanju komunikacijskih izazova i problema interaktivne kooperacije. Časovi ne traju samo 45 minuta, već je njihova dužina direktno srazmerna učeničkom interesovanju za nastavni predmet.



Slika 25. Društvena mreža Edmodo – kreiranje profila

(Izvor: https://www.edmodo.com/?language=sr_RS)

7.7. Wikipedia

Wikipedia na srpskohrvatskom jeziku predstavlja slobodnu otvorenu enciklopediju koju svako može uređivati. Korisnik sam može odrediti da li će pisati ćirilicom ili latinicom. Na Wikipediji postoji oko 200 jezika i broj članaka na srpskohrvatskom jeziku je 427,403. Dokumenti se mogu menjati, i nepravilnosti se mogu prijaviti administratorima. Wikipedia može biti od velike pomoći u edukaciji tako što može da obogati nastavni sadržaj, ali treba uzeti u obzir i to da podaci nisu uvek merodavni jer kao što je navedeno Wikipediju može da uređuje stvara, menja i briše svaki korisnik.

Sadržaj

Prirodne i matematičke nauke/znanosti: astronomija - astrofizika - biologija - geologija - geografija - ekologija - informatika - matematika - medicina - statistika - fizika - hemija

Društvene nauke/znanosti:

antropologija - arheologija - bibliotekarstvo - demografija - istorija - vojne nauke - jezik i lingvistika - kognitivna nauka - muzikologija - obrazovanje - pedagogija - psihologija - sociologija - filozofija - defektologija

Politika, pravo i društvo:

država - društvo - ekonomija - javni poslovi - mediji - političke nauke - pravo - preduzeća - prirodno okruženje - slobodni softver - trgovina - urbanizam

Religija i mitologija:

ateizam - hrišćanstvo/kršćanstvo - islam - judaizam - misticismizam - mitologija - religija - teologija - agnosticizam - rimske pape

Liste:

biografije - godišnji kalendar - dnevni kalendar - internet domeni - države - narodi - kategorije - ljudi - kalendari - liste - članci koje svaka Wikipedia treba imati

Primenjene nauke/znanosti i tehnike:

alat - avijacija - arhitektura - poljoprivreda - komunikacija - energetika - elektronika - industrija - informatika - internet - inženjerstvo - javni poslovi - medicina - slobodni softver - softver - tehnologija - telekomunikacije - transport

Životna svakodnevnica i slobodno vrijeme:

enigmatika - zabava - zdravlje - internet - kuhinja - seksualnost - rekreacija - sport - praznici - turizam - hobi

Umjetnost i kultura:

arhitektura - dizajn - književnost - kultura - mediji - muzika - narodne izreke - ples - teatar - skulptura - slikarstvo - umjetnost - film - grafika

Slika 26. Wikipedija u edukaciji

(Izvor: <https://hr.wikipedia.org/wiki/Obrazovanje>)

Iz svega navedenog se može izvesti zaključak da tradicionalni način odvijanja nastave novim generacijama nije zanimljiv i smatraju ga zastarelim, što stvara nezainteresovanost za gradivo. Korišćenjem novih tehnologija, kao što su društvene mreže, osavremenjuje se sama nastava i učenici postaju sve više zainteresovani za novo gradivo. Treba navesti i druge društvene mreže koje mogu da nađu svoju primenu i u edukaciji, a to su: Instagram, TED, LinkedIn itd.

8. BIG DATA TEHNOLOGIJE POMOĆU KOJIH SE ANALIZIRAJU DRUŠTVENI PODACI

7

U poslednje tri godine ljudsko društvo je stvorilo i sačuvalo više informacija nego u čitavoj istoriji do tada [58].

Kako se ekonomija razvija, tako raste i potreba za korišćenjem podataka intenzivne tehnologije. Trenutno postoji 4,6 milijardi mobilnih telefona širom sveta, a između 1 i 2 milijarde ljudi pristupa internetu preko telefona [59]. Kapacitet razmene informacije putem telekomunikacionih mreža je 281 petabajta u 1986, 471 petabajta u 1993, 2,2 eksabajta 2000, 65 eksabajta u 2007, 667 eksabajta

u 2014, a predviđeno je da će taj broj rapidno da poraste u narednih par godina. Procenjuje se da je jedna trećina od globalno uskladištenih informacija u obliku teksta i alfanumeričkom "Still image" podataka koji je format najkorisniji za većinu Big data tehnologija. Ovo takođe pokazuje potencijal još neiskorišćenih podataka (tj. u obliku video i audio sadržaja).

Prednosti koje donosi tehnologija Big data prepoznale su vodeće softverske kompanije koje isporučuju komercijalni softver. Big data je povećao potražnju stručnjaka za upravljanjem informacija, pa su velike kompanije poput *Software IBM, Microsoft, AG, Oracle Corporation, FICO, SAP, EMC, HP i Dell* uložile više od 15 milijardi dolara u softverske firme specijalizovane za upravljanje podacima i analizu. U 2010. godini ova industrija je bila vredna više od 100 milijardi dolara i raste skoro 10 odsto godišnje, oko dva puta brže od poslovnog softvera u celini [60].

S obzirom da su kompanije uložile veliki novac u istraživanje i razvoj softvera, postoji puno tehnologija koje rade analizu podataka. Neki od najpoznatijih programa za analizu podataka su: Google BigQuery, Amazon SimpleDB, IBM BigInsights, The Revolution Analytics, Ecrion Software, Hypertable, Simply Measured, Oracle-ov Darwin itd. Što se tiče cene ovih programa ona se kreće u rasponima od nekoliko hiljada do nekoliko miliona evra.



Slika 27. Izgled ponude za analizu društvenih mreža kompanije „Simply Measured” [61]

Program koji se činio najpogodnijim za istraživanje Twitter podataka je IBM BigInsights. Verzija koja je korišćena u ovom radu je “IBM InfoSphere BigInsights Quick Start Edition 3,0”, može se besplatno download-ovati sa IBM-ovog oficijalnog sajta. IBM ima i programe koji nude mnogo više od ovoga, ali zato i koštaju puno. Ovo je besplatna verzija programa koji je dovoljan da se preko njega izvrši analiza Twitter podataka koji se mogu primeniti u edukaciji.

U toku izbora alata, najveća konkurencija IBM BigInsights-u je bio alat po imenu Gephi. Gephi predstavlja softver koji je otvorenog koda i koristi se uglavnom za mrežnu analizu i prikaz kompleksnih mreža. Gephi se koristi od strane raznih istraživača, na univerzitetima, u novinarstvu, za prikaz inMaps od strane poslovne društvene mreže LinkedIn itd. Gephi ima razvijen sistem za vizualizaciju pogodan za analizu velikih grafova. Gephi koristi ad-hoc OpenGL engine i kvalitetan softver za vizualizaciju koji može prikazati do 50 000 čvorova i do 500 000 veza. Gephi program

nije besplatan i isključivo je namenjen grafičkom prikazu podataka, tako da je izbor pao na IBM BigInsights koji takođe analizira velike podatke pomoću teorije grafova, ali rezultate prikazuje tabelarno i vrlo je lak za korišćenje.

Sa ciljem da integrišu Hadoop u svoje poslovne softvere, kompanija IBM 2010. godine objavljuje platformu pod nazivom IBM InfoSphere BigInsights. Ovo rešenje daje mogućnost analize velikih raznovrsnih podataka i čvrstu integraciju. Razvijeni su potpuno novi alati BigSheets, Jaql i SystemT pomoću kojih se omogućuje analiziranje u Hadoop okruženju (pored već postojećih open source softvera HBase, Pig, Hive, Avro itd.).

Može se reći da je primena Big data počela pre otprilike deset godina kada su velike kuće Google i Yahoo odlučile da iskoriste velike količina podataka koje su generisale. Sa ciljem da se ti ogromni podaci uspešno obrade 2008. godine je nastao Hadoop kao Apache projekat otvorenog koda. Bitno je pomenuti Map Reduce tehnologiju koja podatke prvenstveno deli u dve faze **mapiranja i redukcije**. U prvoj **fazi mapiranja** podaci **se obrađuju** tako što se dele u uređene parove i mapiraju. Druga faza redukcije sledi posle ovoga i tada se ovi parovi kombinuju u manje skupove i to na kraju budu konačna rešenja. Klasterovanje podata sledi nakon toga i to radi Hadoop tako što podatke u svojim klasterima deli na manje blokove i distribuira ih na NameNode server u okviru klastera gde se kasnije vrši obrada podataka.

Big data tehnologiju čine glavnih pet elemenata: **analitička baza podataka, memorijska baza (In-memory database management system), Hadoop, obrada tokova podataka u realnom vremenu, softver i hardver na kojem se to sve smešta [62].**

Kao podršku svemu ovome u okviru Apache projekta se razvijaju open source projekti koji su vezani za Hadoop. To su: Ambari, Avro, Pig, Spark, HBase, Chukwa, Mahout, Hive i Cassandra. Društvena mreža Twitter koristi Hadoop distribuiranu bazu podataka – Cassandra i ona predstavlja "NoSQL" bazu podataka.

9. SPEKTRALNA TEORIJA GRAFOVA

Analiza društvenih mreža se može opisati kao izučavanje ljudskih odnosa i veza korišćenjem teorije grafova. Društvene mreže se mogu modelovati, simulirati i analizirati pomoću grafova.

Teorija grafova je oblast matematike koja se primenjuje u društvenim mrežama, računarskim mrežama, biomedicinskim mrežama, grid mrežama, transpornim mrežama itd. Grafovi se sastoje od čvorova (koji predstavljaju pojedinačne aktere u okviru mreže) i veza (koje predstavljaju odnose između aktera). Društvena mreža može se opisati skupom iskaza koji opisuju relacije između entiteta mreže. Jednostavan iskaz odnosa dva entiteta predstavlja osnovnu jedinicu analize društvene mreže i označava se kao par. Društvene mreže mogu se opisati grafom $G = (N, E)$, gde je N broj čvorova mreže (eng. nodes), dok E označava broj grana mreže (eng. edges). Svaki par trebalo bi da označi pojedinačni odnos koji se u tradicionalnoj teoriji grafova naziva grana ili veza [63]. Čvorovi individua su povezani sa jednom ili sa više zavisnosti uključujući zajedničke vrednosti, vizije ili ideje, društvene kontakte, konflikte, finansijske razmene. Veze između čvorova obezbeđuju različite tokove, kao npr. poruke, prijateljstvo, bolesti itd.

U analizi društvenih mreža podela na tipove čvorova vrši se na osnovu toga kakva je reprezentacija čvora unutar mreže. Grafovi koji sadrže čvor jednog tipa označavaju se kao jednomodalni, dok se grafovi sa relacijama dva različita tipa čvorova nazivaju bimodalni. Grafovi sa relacijama više različitih tipova nazivaju se multimodalni.

Kao osnovni elementi svake mreže, čvorovi i grane se opisuju dodatnim parametrima koji se dobijaju primenom odgovarajućih algoritama za analizu. Grane u mreži između čvorova mogu biti usmerene (eng. directed) i neusmerene (eng. undirected). Usmerene grane se označavaju kao: $A \rightarrow B$. Neusmerene grane se označavaju sa: $A \leftrightarrow B$ ili $A - B$. Atributi koji bliže opisuju grane su: težina, rangiranje, tip, kao i svojstva u zavisnosti od ostatka strukture mreže.

Metrika čvora predstavlja jednu od osnovnih karakteristika koja se izračunava na osnovu povezanosti svakog čvora u mreži. Stepen čvora se određuje na osnovu povezanosti datog čvora sa ostalim čvorovima u mreži (eng. degree). Vrednost stepena čvora raste ukoliko je povazanost čvora veća. Kod čvorova u usmerenoj mreži razlikuju se ulazni stepen (eng. indegree) pokazuje broj usmerenih grana ka čvoru i izlazni stepen (eng. outdegree) predstavlja broj grana usmerenih od čvora. Kod neusmerenih mreža stepen čvora je jednak ukupnom broju grana koje se stiču u čvoru.

Jedno od bitnih svojstava unutrašnje strukture grafa društvene mreže je centralnost (eng. centrality). Centralnost se opisuje sledećim parametrima:

- Degree centrality (broj najbližih susednih čvorova)
- betweenness centrality (broj najkraćih putanja koje prolaze kroz posmatrani čvor ka ostalim čvorovima u mreži)

- closeness centrality (prosečan broj koraka potrebnih da se od vrha jednog grafa dođe do svih ostalih čvorova)
- Eigenvector centrality (važnost čvora zavisno od važnosti njegovih suseda)
- PageRank centrality (PageRank se može posmatrati kao model ponašanja korisnika, odnosno to je verovatnoća da slučajni surfer poseti određenu stranu na internet)

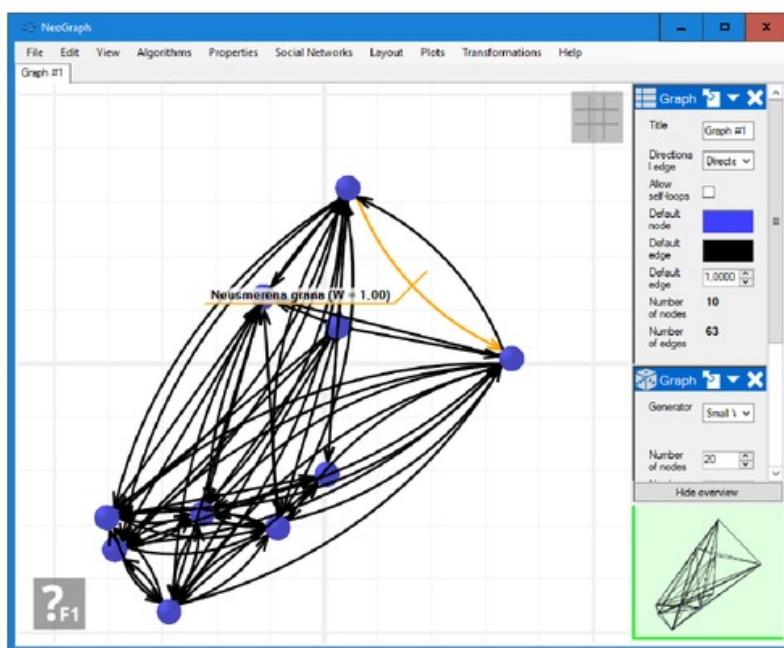
Spektralna teorija grafova je matematička grana gde se susreću linearna algebra i teorija grafova. Za bilo koji dat graf moguće je izvršiti analize spektra grafa bazirane na matičnim predstavama datog grafa. Najvažnije matrice korišćene u analizama spektra grafa su matrica susedstva i Laplacian matrica. Spektralna teorija grafova koristi sopstvene vrednosti (eigenvalues) i prostore sopstvenih vrednosti (eigenspaces) ovih matrica za dobijanje korisnih informacija o mreži, kao što je robusnost, algebarska povezanost itd. [63].

Spektralna teorija grafova se bavi sopstvenim vrednostima i vektorima sopstvenih vrednosti (eigenvectors) grafa iz opšte perspektive. Kod socijalnih mreža čvorovi i grane grafa predstavljaju ljude i relacije između njih. Matrica susedstva je predstava socijalne mreže bazirana na relacijama među ljudima. Laplacian matrica uzima u obzir i stepen čvora. Veoma često običan vektor sopstvenih vrednosti sadrži više informacija o strukturi socijalne zajednice. Na primer Googleov PageRank algoritam koristi spectralnu teoriju grafova za određivanje važnosti veb-sajtova povezanih sa datim upitom pretraživanja. Neke od najviše korišćenih invarijanti grafa su broj čvorova, broj grana, dijametar, radijus, algebarska povezanost, klastering itd. Invarijante (parametri) grafa su snažan analitički alat za istraživanje struktura grafa. Spektar grafa je multiset sopstvenih vrednosti matrice. Spektar grafa je pogodan za pronalaženje informacija o fizičkim karakteristikama mreže.

U radu [67] opisane su i analizirane topološke karakteristike Twitter follow grafa, kao što su raspodela stepena čvora, povezanost komponenata, dužine najkraćih putanja, klastering koeficijenti itd. Twitter mreža ima karakteristike socijalnih mreža kao što su: visok stepen assortativity, male dužine najkraćih putanja, visoki stepeni klasterovanja i visok stepen reciprocity. Ona ima i karakteristike informacionih mreža gde je dominantna interakcija širenje informacija po granama, što je okarakterisano velikim stepenom čvorova, nedostatkom reciprocity i velikim two-hop susedima. Na osnovu analize tih rezultata zaključeno je da je Twitter mreža i informaciona i socijalna mreža. Pošto je follow (praćenje) karakteristika asimetrična, Twitter follow graf je usmeren. Da bi izvršili poređenje raznih karakteristika, u ovom radu su koristili neusmerene mutual (obostran, uzajamni) graf, koji je graf sa recipročnim granama. Jedna grana između dva korisnika u mutual grafu podrazumeva da oba korisnika prate jedan drugog. Od svih grana, 42% u follow grafu

su recipročne grane, tako da je skoro 4 milijarde neusmerenih grana u mutual grafovima. Sve analize u radu [67] su izvršene u Twitter Hadoop analytics stack koristeći Pig softver.

Klastering koeficijent je mera u socijalnim mrežama podskupa korisnika čiji su prijatelji i međusobno prijatelji i njegova vrednost je velika. Na slici 28. je prikazan način prelaska sa usmerenog grafa na neusmereni (mutual) graf gde su dve usmerene grane dva folovera koji prate jedan drugog predstavljene jednom neusmerenom granom (a), kao i dobijena vrednost koeficijenta klasterovanja (b). Korišćeni su podaci sa zvaničnog sajta Stanford Univerziteta za Twitter socijalnu mrežu iz SNAP projekta (<https://snap.stanford.edu/data/>).



a)



b)

Slika 28. Twitter graf

U radovima [64], [65] i [66] prikazani su rezultati istraživanja iz spektralne teorije grafova na neusmerenim grafovima. Ideja je bila da se za pojedine parametre koji su izvedeni iz osnovnih pronadu nove granične vrednosti, definiše fizika stvari i eventualna primena na socijalne mreže.

Neka je G neusmereni povezani graf sa n čvorova i m grana. Ako su $\rho_1 \geq \rho_2 \geq \dots \geq \rho_{n-1} > \rho_n = 0$ normalizovane Laplasove sopstvene vrednosti grafa G , tada se normalizovani linearni i ratio laplasovi spredovi definišu kao

$$NLS(G) = \rho_1 - \rho_{n-1} \text{ i } NRS(G) = \sqrt{\frac{\rho_1}{\rho_{n-1}}} + \sqrt{\frac{\rho_{n-1}}{\rho_1}},$$

respektivno. Dakle, Laplacian spred se definiše kao razlika između najveće i druge najmanje sopstvene vrednosti - algebarske povezanosti (prva najmanja sopstvena vrednost je uvek jednaka nuli) Laplacian matrice grafa G . U radu su određene donje granice za $NLS(G)$ i $NRS(G)$. Ako je broj čvorova n paran i $n \geq 4$, a m broj grana, u radu [64] je dokazano da za normalizovane linearne Laplacian spredove važi nejednačina:

$$NLS(G) = \rho_1 - \rho_{n-1} \geq \sqrt{\frac{2(n-1)R_{-1} - n}{(n-1)^2\alpha(n-1)}}$$

gde je $\alpha(n) = f(x) = \begin{cases} 1/4, & n \text{ parno} \\ (n^2 - 1)/4n^2, & n \text{ neparno} \end{cases}$, a $R_{-1} = \sum_{i \sim j} \frac{1}{d_i d_j}$ Randić indeks.

18

U literaturi postoji veliki broj radova koji se bave optimizacijom mreža koristeći spektralni radijus (maksimalna sopstvena vrednost matrice susedstva), algebarsku povezanost, klasterovanje, i druge parametare iz spektralne teorije grafova. Pojam Laplacian spektralni spred je veoma značajan za opisivanje globalnih karakteristika strukture grafa (npr. diametar). Ovi parametri mogu naći primenu ne samo u algebarskoj teoriji grafova već i u drugim naučnim i tehničkim istraživanjima [64].

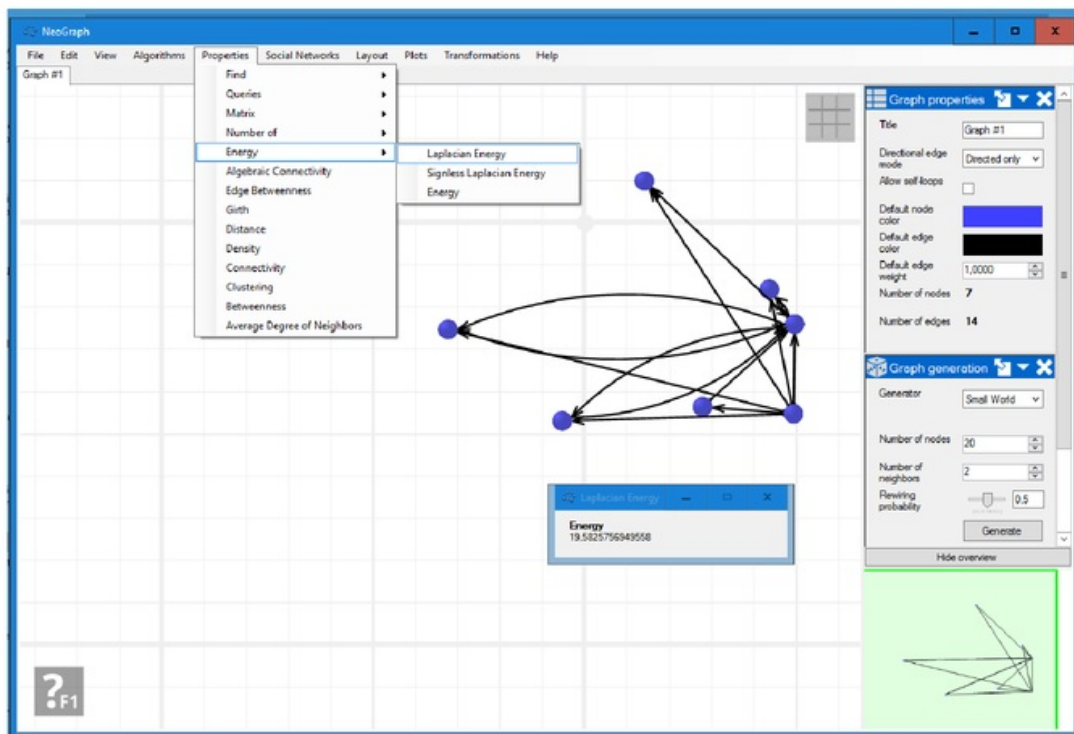
Neka je G neusmereni prost graf sa n čvorova i m grana. Označimo sa $|\lambda|_1 \geq |\lambda|_2 \geq \dots \geq |\lambda|_n$ i $|\rho|_1 \geq |\rho|_2 \geq \dots \geq |\rho|_m$ apsolutne vrednosti sopstvenih vrednosti i Randić sopstvenih vrednosti grafa G poredane u opadajućem redosledu, respektivno. Gornja granica graf invarijante energije $E(G) = \sum_{i=1}^n |\lambda|_i$, a donje i gornje granice invarijante energije $RE(G) = \sum_{i=1}^n |\rho|_i$ su dobijene u radu [65]. U ovom radu je dokazana teorema da je energija grafa $E(G)$ definisana parametrima n , m , $|\lambda|_1$ i $|\lambda|_n$ i da je

$$E(G) \leq \sqrt{2mn - \frac{n}{2} (|\lambda|_1 - |\lambda|_n)^2}$$

kada je n parno a G razdvojena unija $n/2$ putanja K_2 .

U radu [66] za neusmereni povezani graf sa n čvorova i m grana, gde je $n \geq 3$ i gde su μ_i i ρ_i Laplacian i normalizovani Laplacian sopstvene vrednosti, Laplacian-energy-like (LEL) invarijanta grafa G je definisana i određene su donje granice invarijanti $LEL(G)$ predstavljene preko n i m parametara grafa G .

Za parametre koji su istraživani u prethodna tri rada izvršen je veliki broj eksperimenata za Twitter podatke kao i za podatke drugih društvenih mreža (Facebook, Google Plus) koristeći NeoGraph simulator. Pomoću NeoGraph simulatora (razvijen na Matematičkom institutu SANU) može se dobiti prikaz Twitter mreže čiji je graf usmeren i povezan, za koji je određen parametar Laplacian Energy (slika 29).



Slika 29. Primena NeoGraph simulatora na Twitter društvenu mrežu i određivanje Laplacian energije

Pošto je veoma teško dobiti zvanične Twitter podatke za istraživanja na visokoškolskim ustanovama problema edukacije, nastavak istraživanja ove disertacije je krenuo u pravcu korišćenja softvera IBM BigInsights. Korišćeni su podaci dobijeni direktno iz Twitter aplikacije ili iz javno dostupnih baza.

10. REZULTATI ISTRAŽIVANJA

Na osnovu istraživanja iz prethodnih poglavlja odlučeno je da se Twitter društvena mreža analizira pomoću IBM Big data tehnologije. U ovom poglavlju će biti opisani najvažniji rezultati istraživanja. Opisana je IBM BigInsights softverska platforma, proces instalacije i kreiranje fajla za analizu. Potom je opisana analiza i primena Twitter podataka kao i proces izrade aplikacije.

Izabrana je IBM BigInsights softverska platforma zato što može da pomogne u analiziranju velikih količina podataka širokog spektra - podaci koji se često ignorišu ili odbacuju jer su suviše nepraktični ili teški za obradu koristeći tradicionalne metode. Takođe ova platforma nudi mogućnost da se analiziraju Twitter podaci uz pomoć fajla dobijenog pomoću Twitter aplikacije.

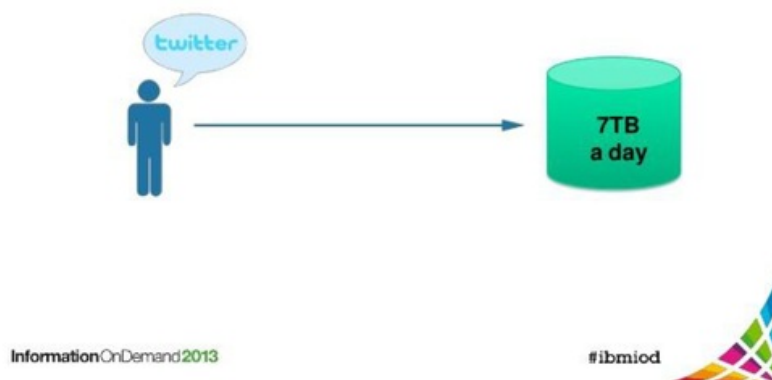
Kako bi pomogli da se efikasno izvedu vrednosti iz takvih podataka, BigInsights je obuhvatio nekoliko open source projekata, uključujući Apache Hadoop i nekoliko tehnologija razvijenih od strane IBM-a, uključujući i BigSheets alat. Hadoop i njemu slični projekti daju efikasan softverski okvir za "data-intensive" aplikacije koje eksploatišu distribuirana kompjuterska okruženja kako bi postigli visok nivo skalabilnosti. IBM tehnologije upotpunjuju ovaj open source okvir sa analitičkim softverom, integrisanim softverom, platformama sa ekstenzijom i alatima.

IBM InfoSphere BigInsights Quick Start Edition je besplatna verzija ove tehnologije koja omogućava nova rešenja kako efikasno pretvoriti velike i složene količine podataka u uvid, kombinovanjem Apache Hadoop (uključujući MapReduce i Hadoop Distributed File Systems) sa jedinstvenim IBM tehnologijama kao što su Big SQL, tekst analiza i BigSheets.

10.1. Instalacija IBM Infosphere Biginsights tehnologije

What is Big Data?

7TB of data processed by Twitter every day



Slika 30. Zanimljiv slajd sa IBM-ove prezentacije o Big data

(Izvor: <https://goo.gl/kMAoEo>)

Za instalaciju IBM InfoSphere BigInsights Quick Start Edition potrebno je zadovoljiti sledeće uslove. Što se tiče hardverskih preduslova za pokretanje i instalaciju, potrebno je minimum 4GB RAM memorije, 40GB prostora na hard disku i dva procesora. Softver, tj operativni sistem na kojem radi BigInsight je Linux. Da bi se uspešno pokrenula instalacija IBM BigInsights, koja zahteva Linux operativni sistem, potrebno je u par koraka izvršiti pripremu.

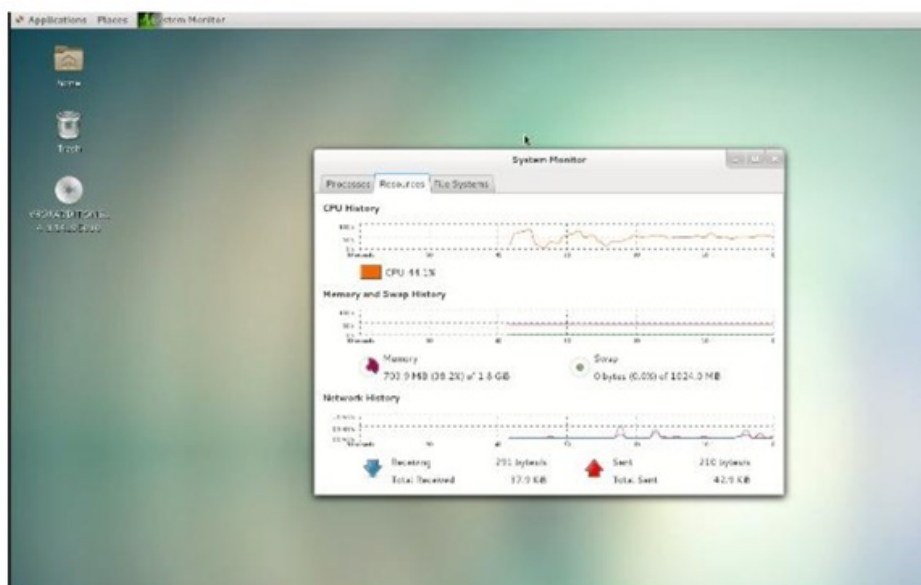
- Prvi korak - instalacija VMware Workstation

S obzirom da većina računara koristi Windows operativni sistem, potrebno je instalirati virtuelnu mašinu na kojoj će se instalirati Linux a potom i IBM BigInsight. Pre instalacije virtualne mašine obavezno treba proveriti da li tip procesora podržava virtuelizaciju. Treba iz BIOS-a otići u „ADVANCE MODE“ i aktivirati opciju „Intel virtualization technology“ koja je uglavnom fabrički isključena. Za ovaj deo instalacije korišćen je jedan od najpoznatijih softvera namenjen virtuelizaciji - VMware Workstation. To je softver koji pruža mogućnost korisniku da kreira neograničen broj virtuelnih mašina na računaru, gde svaka virtuelna mašina može imati svoj „gostujući“ operativni sistem. VMware Workstation dopušta jednoj fizičkoj mašini da pokrene neograničeno mnogo OS istovremeno. Ostali VMware produkti pomažu rukovođenje (VMware

Player) i prenošenje (VMware Server) sa određenim prednostima ali i manama u odnosu na VMware Workstation.

- Drugi korak - instalacija CentOS

U drugom delu instalacije potrebno je instalirati operativni sistem Linux koji se svakodnevno koristi u raznim uređajima, od mobilnih telefona do televizora. S obzirom na to koliko je Linux rasprostranjen danas postoji i veliki broj distribucija kao što je CentOS (Community Enterprise Operating System). CentOS je lak je za korišćenje i podržava IBM BigInsight. Na samom početku instalacije, VMware već nudi opciju CentOS, tako da instalacija može da počne bez ikakvih poteškoća.



Slika 31. Izgled CentOS nakon uspešne instalacije

- Treći korak - IBM BigInsights

Nakon uspešnog instaliranja novog operativnog sistema CentOS, potrebno je logovati se sa „root“ nalogom, da bi se pokrenula instalacija koju je moguće preuzeti sa sajta IBM-a koji nudi najnoviju verziju IBM InfoSphere BigInsights Quick Start Edition V3.0 [68]. Za početak je u terminalu potrebno upisati par naredbi koje idu sledećim redom:

1. Create user biadmin
2. Add biadmin in visudo
3. Login as biadmin
4. Create sudo mkdir /hadoop and change the ownership (sudo chown -R biadmin:biadmin /hadoop)
 Create sudo mkdir /opt/biadmin and change the ownership (sudo chown -R biadmin:biadmin /opt/biadmin)
 Create sudo mkdir -p /var/ibm/biginsights and change ownership (sudo chown -R biadmin:biadmin /var/ibm/)
 Create sudo mkdir /opt/ibm and change the ownership (sudo chown -R biadmin:biadmin /opt/ibm/)
5. Create passwordless login
 ssh-keygen
 ssh-copy-id -i ~/.ssh/id_rsa.pub localhost
 ssh localhost (to check)
6. Mount the shared folder (sudo mount -t vboxsf Downloads Desktop/sharing)
7. Extract Biginsights files (tar -xvzf Desktop/sharing/)
8. Start the Biginsights installer (./start-all.sh)

Kada se otpakuju BigInsights fajlovi sa desktopa, pokreće se komanda start i instalacija IBM BigInsight počinje.

```

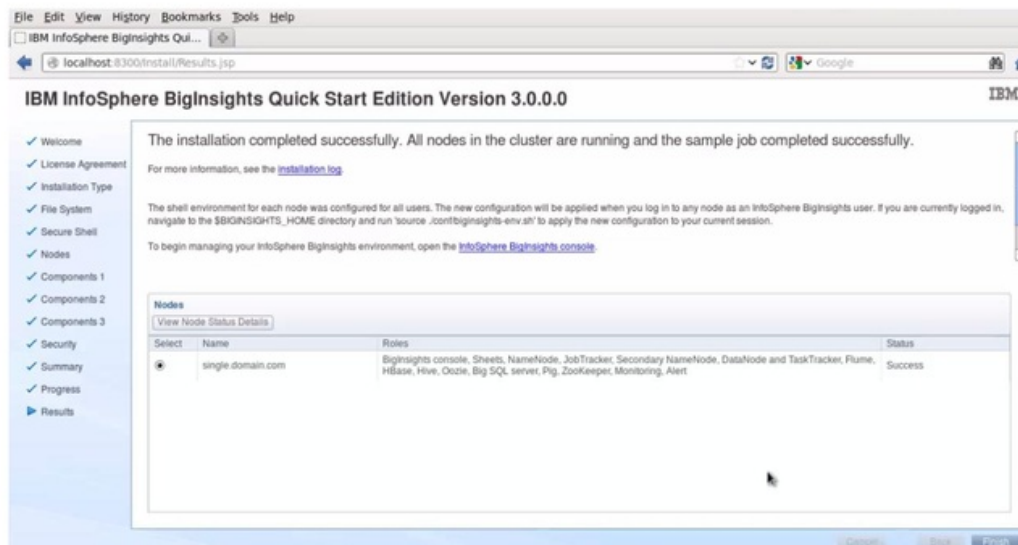
biadmin@localhost:~/biginsights-basic-linux64_b20120406_0812
File Edit View Terminal Tabs Help
_b20120406_0812/installer-console/_version.properties
[biadmin@localhost ~]$ ls
biginsights-basic-linux64_b20120406_0812 Desktop
[biadmin@localhost ~]$ clear

[biadmin@localhost ~]$ ls
biginsights-basic-linux64_b20120406_0812 Desktop
[biadmin@localhost ~]$ cd biginsights-basic-linux64_b20120406_0812/
[biadmin@localhost biginsights-basic-linux64_b20120406_0812]$ ls
artifacts installer-console README.html silent-install
installer licenses README.txt start.sh
[biadmin@localhost biginsights-basic-linux64_b20120406_0812]$ ./start.sh
Extracting Java ....
Java extraction complete, using JAVA_HOME=/home/biadmin/biginsights-basic-linux6
4_b20120406_0812/_jvm/ibm-java-x86_64-60
Verifying port 8300 availability
port 8300 available
Starting BigInsights Installer .....|

```

Slika 32. Startovanje instalacije IBM BigInsights

Ako su sva prethodna podešavanja bila u redu, potrebno je otvoriti Mozillu i ukucati adresu „localhost:8300/Install“. U tom delu se konfigurišu opcije koje su vezane za tip instalacije, sigurnost, licencu itd.



Slika 33. Prikaz završene instalacije IBM BigInsights

10.2. Proces kreiranja fajla za analizu

Za uspešnu analizu društvene mreže Twitter, potrebno je kroz nekoliko koraka doći do potrebnih rezultata. U ovom delu rada će se detaljno opisati način na koji se analiziraju podaci i postupak pripreme potrebnih fajlova koje će kasnije IBM BigInsights iskoristiti u analizi i virtuelizaciji.

Twitter održava otvorenu platformu koja podržava milione ljudi širom sveta koji dele i otkrivaju šta se dešava. Twitter želi da osnaži svoje partnere i korisnike, da bi se izgradila vredna mreža informacija, time što dozvoljava besplatno kreiranje raznih aplikacija vezanih za Twitter analizu. Za početak je potrebno otići na sajt za Twitter developere [69] i ulogovati se preko Twitter naloga. Proces kreiranja aplikacije se realizuje u nekoliko koraka (<https://apps.Twitter.com/>) pri čemu se definišu: naziv aplikacije, podešavanja aplikacije, pristupni tokeni, dozvole za rad aplikacije itd. Ukoliko se ne poseduje nalog, potrebno je otvoriti ga na Twitter sajtu. Sledeći korak je kreiranje aplikacije, gde je potrebno uneti lične podatke i lokacije gde će se smeštati fajlovi za analizu. Takođe se na kraju traži upit preko kog će Twitter znati šta je potrebno pretražiti. U ovom primeru je pretraga vezana za „android“ i izgleda ovako:

q=android&count=1000

Uz pomoć ovog upita, Twitter će zadnjih 1000 tviteva, koji sadrže u sebi reč „android“, da generiše u kod koji je potreban u sledećem koraku da bi se napravio fajl iz koga će se kasnije analizirati podaci.

Request Settings

Request type: *

☒ GET
☐ POST
☐ DELETE
☐ PUT
☐ HEAD

Request URI: *

The full URI, without parameters. For example: `https://api.twitter.com/1.1/statuses/home_timeline.json`

Request query:

The parameters for your request. For example: `include_entities=true&page=2`. Note these parameters will be sent on the querystring for GET requests, and in the request body for POST requests.

[See OAuth signature for this request](#)

Slika 34. Kreiranje zahteva za analizu u Twitter aplikaciji

Nakon toga se dobija komanda pod nazivom „cURL comand“ koju je potrebno iskoristiti u roku od par minuta jer nakon toga više neće važiti. To je obaveštenje koje se dobije nakon što se generiše kod iz upita tj. zahteva koji je postavljen u prvom koraku.

OAuth Signing Results

Important: This will only be valid for a few minutes. Also remember the cURL command will actually execute the request.

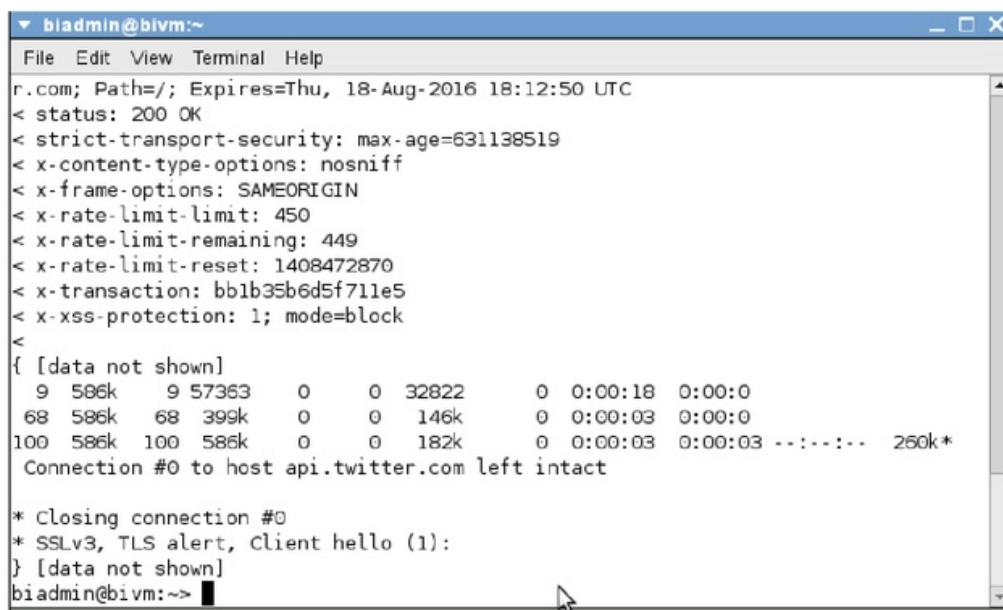
Signature base string	GEThttps%3A%2F%2Fapi.twitter.com%2F1.1%2Fsearch%2Ftweets.json&count%3D1000%26oauth_consumer_key%3DB6tJcNpZDwXM3DiFDkMXKbthE%26oauth_nonce%3Dfba1269c6ce649e37b571ac57044c59b%26oauth_signature_method%3DHMAC-SHA1%26oauth_timestamp%3D1408404534%26oauth_token%3D%26oauth_version%3D1.0%26q%3Dandroid
Authorization header	Authorization: OAuth oauth_consumer_key="B6tJcNpZDwXM3DiFDkMXKbthE", oauth_nonce="fba1269c6ce649e37b571ac57044c59b", oauth_signature="BRYu8hAgjwDoAPqagCalchQJ9Ww%3D", oauth_signature_method="HMAC-SHA1", oauth_timestamp="1408404534", oauth_token="", oauth_version="1.0"
cURL command	curl --get 'https://api.twitter.com/1.1/search/tweets.json' --data 'count=1000&q=android' --header 'Authorization: OAuth oauth_consumer_key="B6tJcNpZDwXM3DiFDkMXKbthE", oauth_nonce="fba1269c6ce649e37b571ac57044c59b", oauth_signature="BRYu8hAgjwDoAPqagCalchQJ9Ww%3D", oauth_signature_method="HMAC-SHA1", oauth_timestamp="1408404534", oauth_token="", oauth_version="1.0" --verbose

Slika 35. cURL komanda potrebna za kreiranje „json“ fajla

Potrebno je kopirati „cURL command“ i otvoriti terminal u kom će se zalepiti ceo kod, a nakon toga treba odrediti ime fajla sa ekstenzijom „json“. U nastavku se ukuca ime fajla koji će da se kreira, npr „>>and.json“, i pokreće se komanda koja je prikazana na slici 36.



Slika 36. Pokretanje cURL command-e



```
biadmin@bivm:~  
File Edit View Terminal Help  
r.com; Path=/; Expires=Thu, 18-Aug-2016 18:12:50 UTC  
< status: 200 OK  
< strict-transport-security: max-age=631138519  
< x-content-type-options: nosniff  
< x-frame-options: SAMEORIGIN  
< x-rate-limit-limit: 450  
< x-rate-limit-remaining: 449  
< x-rate-limit-reset: 1408472870  
< x-transaction: bblb35b6d5f711e5  
< x-xss-protection: 1; mode=block  
<  
{ [data not shown]  
  9 586k  9 57363  0  0 32822  0 0:00:18 0:00:0  
 68 586k  68 399k  0  0 146k  0 0:00:03 0:00:0  
100 586k 100 586k  0  0 182k  0 0:00:03 0:00:03 --:--:-- 260k*  
Connection #0 to host api.twitter.com left intact  
  
* Closing connection #0  
* SSLv3, TLS alert, Client hello (1):  
} [data not shown]  
biadmin@bivm:~>
```

Slika 37. Izgled terminala nakon završetka

Posle ovoga, u biadmin dokumentima nalaziće se fajl pod imenom „and.json“. "JavaScript Object Notation" (JSON) predstavlja standard koji je jednostavan, tekstualan i otvoren. Dizajniran je za čitljivu razmenu podataka i izveden je iz JavaScript jezika za predstavljanje jednostavnih struktura podataka i asocijativnih nizova, odnosno objekata. I pored njegove bliske veze sa JavaScript-om, JSON je nezavisan jezik. Često se koristi za serijalizaciju i prenos struktuiranih podataka preko mrežne veze, prvenstveno između servera i veb aplikacija, služeći kao alternativa XML formatu.

Obrada i transformacija podataka se vrši sa JAQL. JAQL je prvenstveno jezik upita za "JavaScript Object Notation", ali ne podržava samo JSON, već i XML, CSV itd. On omogućava obradu i strukturiranih i netradicionalnih podataka i doniran je od strane IBM za open source zajednice. Konkretno, JAQL omogućava da se izaberu, pridruže, grupišu, i filtriraju podaci koji su uskladišteni u HDFS, slično kao spoj Pig i Hive. To je upitni jezik inspirisan mnogim programskim i upitnim jezicima, uključujući i Lisp, SQL, XQuery, i Pig. JAQL je funkcionalan, deklarativni upitni jezik koji je dizajniran za obradu velikih skupova podataka, takode i prerađuje upite na visokom nivou, i kada je potrebno u "niskom nivou" koji se sastoji od upita MapReduce poslova. Sada su tвитови prikupljeni u JSON formatu koji se čuva u /home/biadmin/Documents/and.json. Sledeća stvar koja je potrebna da se uradi je da se tвитови pretvore u jednostavnije strukture koje se skladište u HDFS kao da su razdvojeni zarezima. Za to će se koristiti JAQL koji se već nalazi u IBM BigInsights.

Potreban je JAQL plugin za Eclipse koji takođe nije podržan za Windows operativni sistem, a koji je dostupan već na stranici dobrodošlice kada se instalira IBM BigInsights.

Eclipse predstavlja programsku razvojnu okolinu pisanu u Javi. Najčešće se koristiti pri razvoju aplikacija u Javi, PHP, Python, R, Ruby, Android, Ada, C, C++, COBOL, Perl, Scala, Clojure, Scheme itd. Pre instalacije Eclipse neophodno je instalirati odgovarajuću verziju JDK-a.

Potrebno je otići na veb konzolu (<http://localhost:8080/>) i kliknuti na dugme "Omogućite Eclipse razvojno okruženje za razvoj aplikacija BigInsights". Zatim treba download-ovati Eclipse 3.6 koja je jedina podržana verzija u BigInsights. Nakon startovanja Eclipse, potrebno je izvršiti neka od podešavanja od kojih su najbitnija- povezivanje sa BigInsight serverom, i kreiranje novog projekta gde se određuje lokacija fajlova koji su potrebni za konvertovanje, i nakon toga upisati kod JAQL koji učitava tvitove i snima rezultate na HDFS. Zatim pravi dva fajla, ".dell" - zarezima odvojene reči, on će nadalje biti potreban, i "seq" koji sadži samo tekstualni sadržaj tvitova.

- 1) jaqlGet() - učitava JSON tekstualni fajl
- 2) tweets.statuses -> transform ... - iz original JSON fajla transformiše ga u strukturirani prosti fajl
- 3) tweets -> write(del("/user/root/tweets.del")) - pravi novi strukturirani, zarezom odvojeni fajl koji će se kasnije koristiti za analizu.
- 4) tweets -> transform \$.text; - sadrži samo tekstualni deo tvita
- 3
- 5) tweets_text -> write(seq("/user/root/content/tweets_text.seq")); - skladišti tekstualni sadržaj tvita u Hadoop fajl.

Kod u Eclipse treba da izleda ovako, sa pretragom za „android“ - primer korišćen i u prethodnom koraku.

```
//TODO: Add JAQL content here
3
tweets = jaqlGet("file:///home/biadmin/and.json");

tweets = tweets.statuses -> transform {

    $.created_at,

    tweet_id: $.id_str,

    $.geo,

    $.coordinates,
```

```

    $.location,

    user_id: $.user.id_str,

    user_name: $.user.name,

    user_screen_name: $.user.screen_name,

    user_location: $.user.location,

    user_description: $.user.description,

    user_url: $.user.url,

    user_followers_count: $.user.followers_count,

    user_friends_count: $.user.friends_count,

    $.retweet_count,

    $.favorite_count,

    $.lang,

    text: $.text };

3
tweets -> write(del("file:///home/biadmin/and.del", schema = schema {

    created_at,

    tweet_id,

    geo,

    coordinates,

    location,

    user_id,

    user_name,

    user_screen_name,

    user_location,

    user_description,

    user_url,

    user_followers_count,

    user_friends_count,

```

```
retweet_count,
```

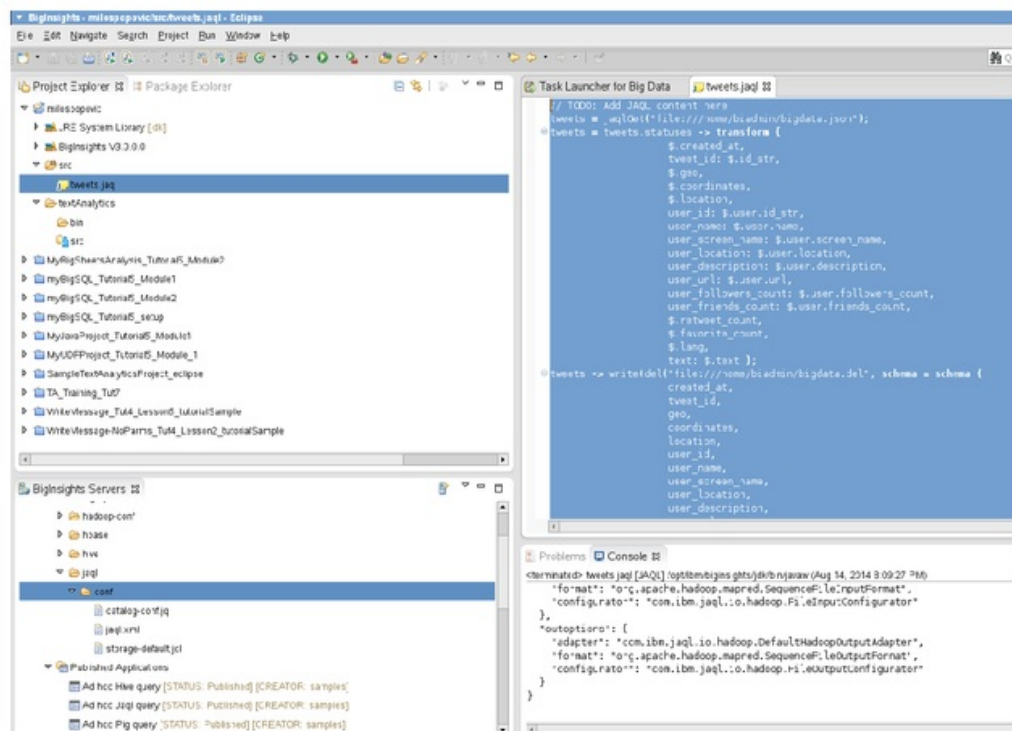
```
favorite_count,
```

```
lang,
```

```
text }));
```

```
tweets_text = tweets -> transform $.text;
```

```
tweets_text -> write(seq("file:///home/biadmin/content/and_text.seq"));
```

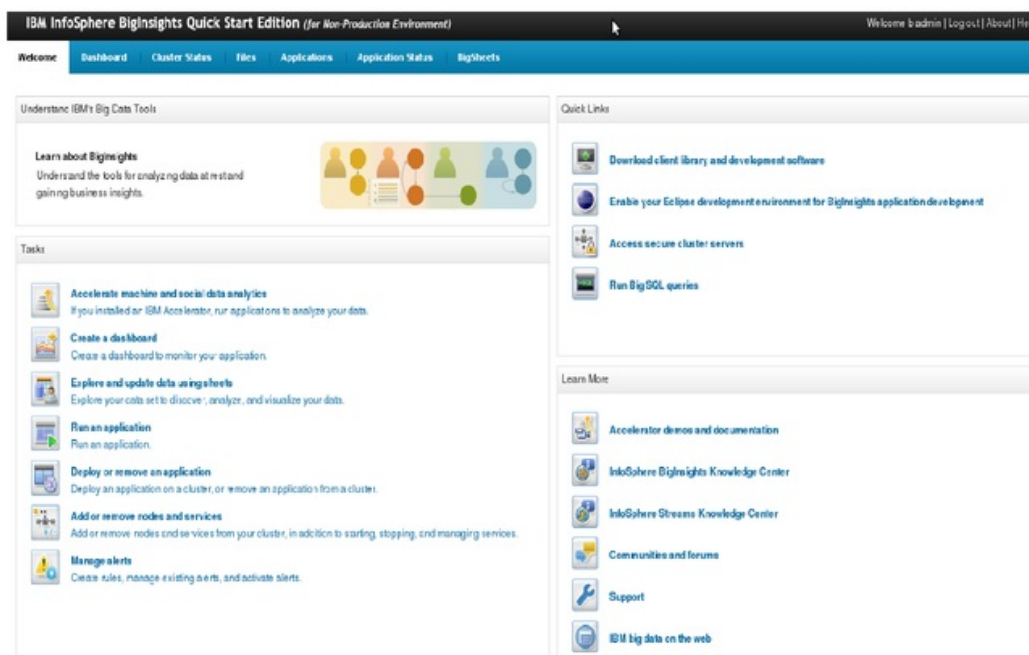


Slika 38. Izgled Eclipse prozora nakon uspješno izvršenih podešavanja

I na kraju je potrebno pokrenuti Eclipse tako što se klikne na „Run the JAQL statement“ i pokreće se kod koji u roku od par sekundi kreira fajl „and.dell“ koji je potreban za analizu u IBM BigInsights.

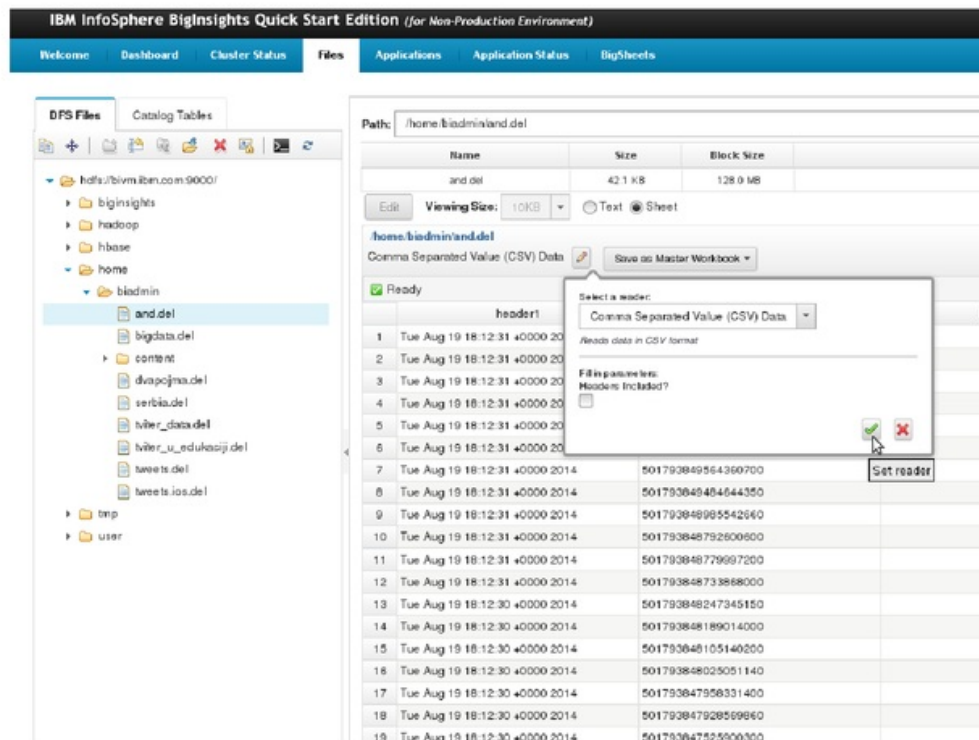
10.3. Analiza Twitter podataka

Nakon što je kreiran fajl „and.dell“ sa sadržajem od hiljadu poslednjih tvitova koji u sebi sadrže pojam „android“, potrebno je pokrenuti IBM InfoSphere BigInsights Quick Start Edition tako što se otvori Veb konzola i ukuca se adresa <http://bivm:8080>.



Slika 39. Početna stranica IBM BigInsights

U sledećem koraku potrebno je pronaći fajl koji je kreiran - „and.dell“ i upload-ovati ga u tabu „Files“ gde će se automatski prepoznati kao tekstualni fajl sa odlikama: Veličina 42KB, Blok veličina 128MB. Da bi nadalje tekst bio vidljiv potrebno je izabrati opciju Sheet, da bi se podaci sortirali po kolonama i redovima, i treba odabrati drugi format za čitanje linija, a to je tekstualni format CSV gde su zarezom razdvojene vrednosti.



Slika 40. Učitavanje „and.del“ fajla u CSV formatu

U daljem postupku treba snimiti fajl kao master radnu svesku pod imenom „analiza androida“, koja se posle toga automatski otvara u BigSheets alatu. BigSheets je alat za analizu baziran na pretraživaču koji je prvobitno razvila grupa novih tehnologija IBM-a (Emerging Technologies group). Danas je BigSheets uključen u BigInsights sa ciljem da se korisnicima omogući istraživanje i analiza podataka. BigSheets predstavlja tabelarni interfejs, tako da korisnici mogu da kreiraju, filtriraju, kombinuju, istražuju, i šematski prikazuju podatke prikupljene iz različitih izvora. BigInsights veb konzola uključuje karticu (tab) na vrhu za pristup BigSheets-u. BigSheets je alat za analitičare obezbeđen od strane IBM InfoSphere® BigInsights, i predstavlja platformu zasnovanu na open source Apache Hadoop projektu. BigSheets je tabelarni alat koji se uglavnom koristi za analizu društvenih medija i strukturiranih podataka prikupljenih kroz aplikacije uzoraka dobijenih iz BigInsights-a. Kada BigInsights prikupi i ubaci podatke iz različitih izvora, BigSheets alat omogućava interaktivna istraživanja kao i rukovanje sa podacima iz distribuiranog sistema podataka. U BigSheets alatu se lako formiraju podaci i njima se rukuje pomoću ugrađenih macro funkcija. Takođe je moguće napraviti grafikone za vizualni prikaz rada kao i eksportovati rezultate analize u jednom od nekoliko popularnih izlaznih formata.

Kada se otvori radna sveska u alatu BigSheet, mogu se videti podaci za koje je izvršen upit. Tj 1000 poslednjih tvitova koji sadrže pojam “android”. Sadržaj i struktura ove radne sveske “analiza android” se ne može menjati, i ona predstavlja glavni izvor podataka. IBM BigInsights je ovaj fajl sortirao po kolonama, ima ih 17 i njihovo značenje je predstavljeno u tabeli 5.

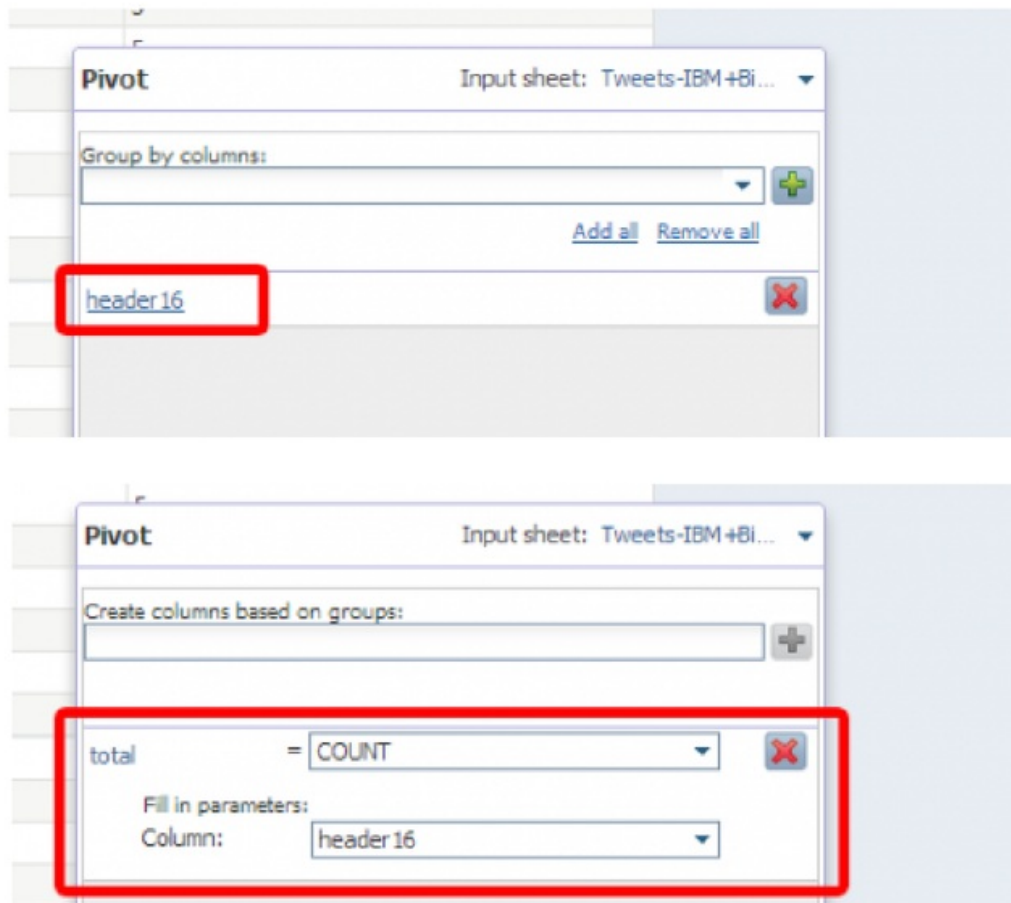
header1	(eng - created_at) - Datum i vreme kada je tvit kreiran
header2	(eng - id_str) - ID broj
header3	(eng - geo) - Geografski položaj
header4	(eng - coordinates) - Koordinate
header5	(eng - location) - Lokacija
header6	(eng - user.id_str) - ID korisnika
header7	(eng - user.name) - Korisničko ime
header8	(eng - user.screen_name) - Korisničko Twitter ime
header9	(eng - user.location) - Lokacija korisnika
header10	(eng - user.description) - Kratko objašnjenje o korisniku
header11	(eng - user.url) - Link ka korisniku
header12	(eng - user.followers_count) - Broj pratilaca
header13	(eng - user.friends_count) - Broj “prijatelja” koje korisnik prati

header14	(eng - retweet_count) - Broj "retvita"
header15	(eng - favorite_count) – Broj "omiljenih"
header16	(eng - lang) - Jezik kojim se korisnik služi
header17	(eng - text) - Kompletan tweet u vidu teksta

Tabela 5. Značenje kolona u IBM BigInsights izveštaju

Na osnovu ovih podataka moguće je izvršiti nekoliko tipova virtualizacije. Na primer - Pokrivenost jezika u grafikonu (pita). Grafikon treba da pokazuje ukupan procenat pokrivenosti jezika - koji se jezik najviše koristio u pisanju tweetova koji se analiziraju. Da bi se postigao ovaj rezultat, potrebno je kreirati grupu koja će sadržati kolonu header16 (Lang - jezik) i obezbedi ukupan broj tweetova u svakoj grupi. Da bi se to uradilo potrebno je kliknuti na "Build new workbook", jer kao što je već pomenuto, master radna sveska se ne može modifikovati, tako da je potrebno napraviti novu radnu svesku na osnovu master radne sveske "analiza android".

Kada se otvori nova radna sveska, klikne se na dugme "Add sheets", pa na opciju "Group" gde je potrebno izabrati kolonu header 16 tj. kolonu koja prikazuje kojim jezikom je napisan tweet. Sledeći korak je opcija COUNT gde je potrebno kreirati kolonu Total, koja će pokrenuti funkciju COUNT za izračunavanje broja istih jezika tj. pojmova iz kolone header16. Primer je prikazan na slici 41.



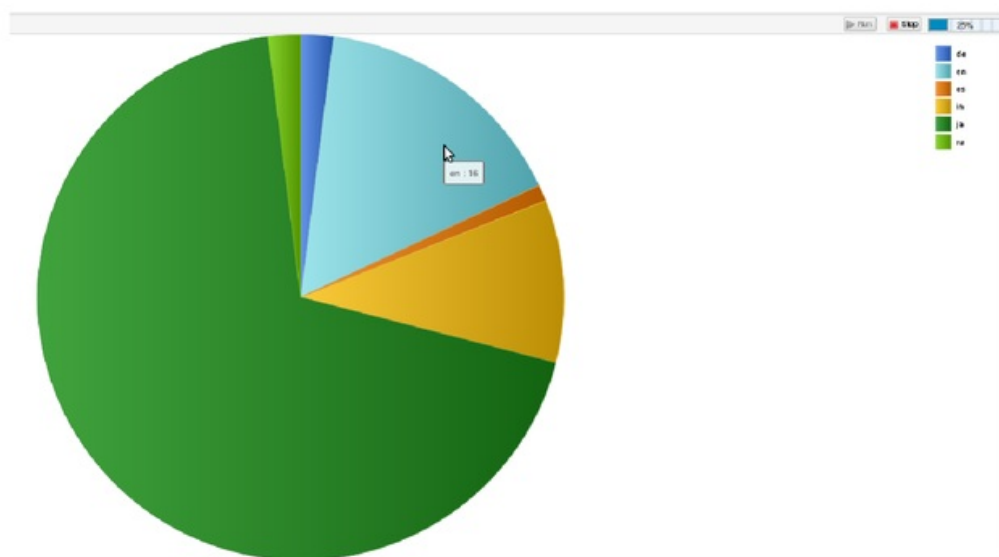
Slika 41. Funkcija count

Posle ove grupacije, dobija se nova radna sveska koja sadrži dve kolone i izgleda kao na slici 42.

	A	B
	header16	total
1	de	2
2	en	16
3	es	1
4	in	10
5	ja	69
6	ru	2

Slika 42. Izgled dobijene tabele

Za vizualizaciju je potrebno kliknuti na “Add chart” u donjem levom uglu i izabrati tip grafikona koji će prikazivati kojim je jezicima najviše tvitova napisano koji su vezani za android. Iza svega ovoga, BigSheets izvršava Pig skripte koje iniciraju MapReduce posao. Svaki put kada se promeni izvor podataka, potrebno je ponovo pokrenuti grafikon i dobijaju se sveži rezultati. U ovom primeru je naveden grafikon tipa “pie - astro”.



Slika 43. Vizuelni prikaz kojim jezikom se najviše tvitovalo za pojam “android”

Iz ovog grafikona se može videti da u zadnjih 1000 tviteva, skoro dve trećine je napisano na Japanskom jeziku, sledeći je Engleski, pa onda Nemački itd.

Postoje tri osnovne vrste vizualizacije: MAP, CLOUD, CHART. Map vizualizacija može da predstavi podatke na mapi sveta. Cloud vizualizacija predstavlja vrednosti tako što uveliča reč koja ima najveću vrednost, i prelaskom mišem preko reči vidi se broj koji je potreban. Relativna veličina reči odražava njenu vrednost, na primer, analiza novog elektronskog uređaja na medijskom sajtu društvene mreže, može doneti sledeće reči na osnovu učestalosti kojom se pojavljuju na sajtu: jednostavan, brz, skupo itd. Chart vizualizacija je klasična i podrazumeva korišćenje jednog od nekoliko tipova grafikona da bi se grafički prikazali podaci.

10.4. Primena podataka u edukaciji

U ovom delu rada je opisano kako radi Hadoop ekosistem sa primerom kako je moguće iskoristiti Twitter podatke za poboljšanje kvaliteta nastave na nekim predmetima u Beogradskoj poslovnoj školi.

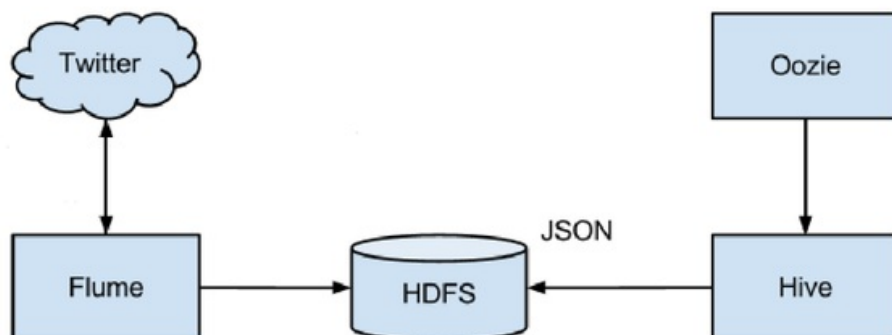
Studenti se mogu upoznati sa tehnikama za analizu društvenih mreža, gde mogu uvideti primenu različitih tipova društvenih medija i kako se društvene mreže mogu koristiti kao alat primenjen kroz interakciju.

Prvi primer opisuje kako je moguće Twitter podatke koristiti na smeru “marketing”.

Društveni mediji su stekli veliku popularnost kod marketinških timova, a Twitter predstavlja efikasno sredstvo za jednu kompaniju pomoću koje skreće pažnju na svoje proizvode. Twitter olakšava angažovanje korisnika i direktnu komunikaciju sa njima, a zauzvrat korisnici mogu kompaniji da obezbede marketing “od usta do usta” tako što će pričati o njihovim proizvodima. Pored toga što je analizom određenih pojmova u marketingu moguće uvideti reakciju korisnika na konkretan proizvod, moguće je otkriti i neku novu ideju za reklamiranje. Marketinški sektori mogu biti efikasniji tako što će selektivno birati do koga žele da dopru, a to mogu da urade tako što otkriju ko je najviše uticajan na Twitteru.

Kada određeni korisnik ažurira svoj Twitter status, to ažuriranje vide svi njegovi pratioci. Korisnik takođe može da “retvituje” tvitove drugih korisnika. Ritvitovanje predstavlja proces sličan prosleđivanju nekog mejla. Ukoliko korisnik 1 vidi tvit od korisnika 2, i ritvituje ga, svi pratioci korisnika 1 će videti tvit od korisnika 2, čak i ako ga ne prate na Twitteru. Kroz ritvitove poruke mogu otići mnogo dalje od samih pratioca osobe koja je prvobitno poslala tvit. Imajući u vidu tu činjenicu potrebno je pokušati da se angažuju korisnici čije objave imaju tendenciju da generišu mnogo ritvitova. Budući da Twitter beleži broj svih ritvitova, analizom Twitter podataka, moguće je doći do korisnika koji su ciljna grupa. Potrebno je analizom podataka doći do odgovora na pitanja - koji Twitter korisnici imaju najviše ritvitova i ko su najuticajni ljudi u “industriji”.

Ispitivanje Twitter podataka u tradicionalnoj RDBMS je nezgodno, budući da Twitter Streaming API izbacuje tvitove u JSON formatu koji može biti proizvoljno složen. U Hadoop ekosistemu, projekat Hive omogućava interfejs upita koji se može iskoristiti za ispitivanje podataka koji se nalaze u HDFS-u. Jezik ovog upita veoma podseća na SQL, ali omogućava da se sa lakoćom oblikuju složeni tipovi, tako da je lako ispitati tip podataka koji se poseduju. Dakle, dobijeni Twitter podaci se stavljaju u Hive tako što se najpre prebace u HDFS.



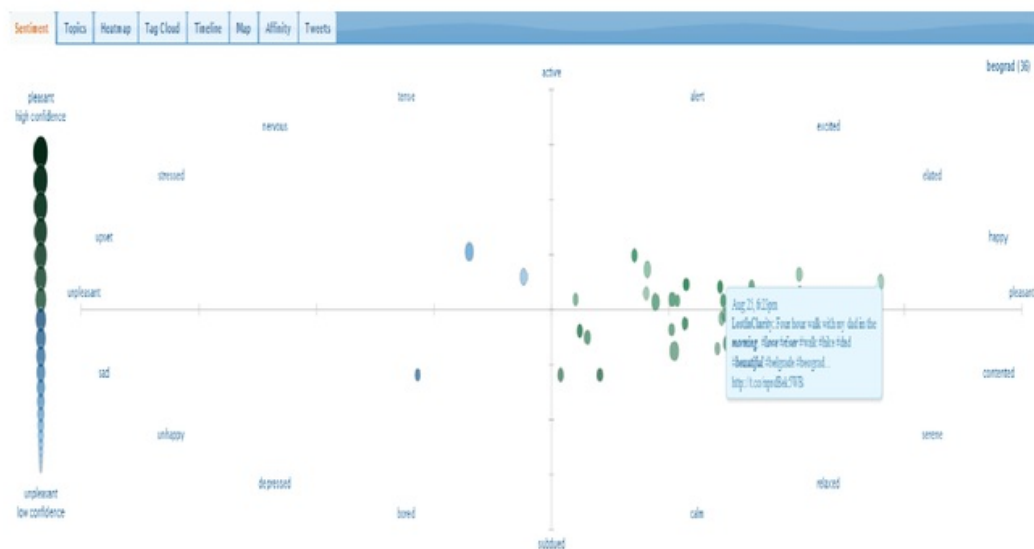
Slika 44. Prikaz kako se neke komponente mogu složiti [70]

Prikupljanje podataka se vrši pomoću Apache Flume koji predstavlja sistem za unos podataka koji je konfigurisan definisanjem krajnjih tačaka u protoku podataka koji se nazivaju "source" i "sink". U Flume-u, svaki pojedinačni deo podataka (tvitova) se naziva događaj; "source" proizvodi događaje, šalje događaje kroz kanal koji povezuje "source" i "sink". Sink zatim prosleđuje događaje prethodno definisanoj lokaciji - piše HDFS fajlovima. Flume podržava neke standardne izvore podataka, kao što su syslog ili netcat. Moguće je koristiti prilagođeni "source" kako bi filtrirali tvitove na bazi ključnih reči što bi pomoglo u identifikaciji relevantnih tvitova umesto čistog uzorka svih Twitter podataka. Pre nego što se ispituju podaci mora se biti siguran da Hive tabela može pravilno da tumači JSON podatke. Uobičajno, Hive očekuje da ulazni fajlovi koriste određeni format redova ali Twitter podaci su u JSON formatu, koji neće raditi po unapred određenim principima. To je zapravo jedna od najvećih prednosti Hive-a. Hive omogućava da se fleksibilno definiše i redefiniše način na koji su podaci predstavljeni na disku. Takođe Hive će prevesti JSON strukture u kolone koje se ispituju, a kao rezultat se dobijaju sortirani korisnici po broju ritviteva. Nakon ovoga se uspešno sastavlja detaljan sistem, koji prikuplja podatke iz Twitter Streaming API-ja, šalje tvitove fajlovima na HDFS kroz Flume, i koristi Oozie za periodično učitavanje fajlova u Hive-u, gde je moguće ispitati neobrađene JSON podatke. Iz ovih rezultata se može videti čiji tvitovi dopiru do najšire publike, ali i utvrditi da li ovi ljudi redovno komuniciraju ili ne. Ove informacije se mogu iskoristiti da se pažljivije usmere poruke kako bi korisnici pričali o proizvodima, a zauzvrat drugi ljudi će dalje širiti priču o proizvodima. Ovakva konstrukcija se može iskoristiti za razne aplikacije dizajnirane za pregledanje Twitter podataka, kao što su one za

prepoznavanje spem naloga, ili one za identifikovanje grupe ključnih reči. Uključivanjem različitih Flume izvora i Hive-a, aplikacija može biti prilagođena za na primer analizu veb logova itd.

Hadoop klasteri sa visokim performansama u IBM BigInsights-u su veoma skupi. U ovom istraživanju je korišćena besplatna verzija ove tehnologije koja ne daje korisniku mogućnost da koristi naprednije opcije. U ovom delu rada je dat primer kako se mogu u realnom vremenu (ili sa veoma malim zakašnjenjem) obrađivati veliki podaci koji se trenutno prikupljaju. U naprednoj verziji ova tehnologije omogućava otkrivanje i sumiranje ukupnog osećanja. Moguće je izvršiti sentimentalnu analizu Twitter podataka koja bi pomogla u poboljšanju kvaliteta nastave na nekim predmetima na smeru “turizam”.

Mikroblogovi su evoluirali i postali izvor različitih vrsta informacija. To se dešava zbog prirode mikroblogova na kojima ljudi postavljaju u realnom vremenu poruke o njihovim mišljenjima o različitim temama, raspravljaju o aktuelnim pitanjima, žale se ili izražavaju pozitivna osećanja za destinacije koje obilaze u svakodnevnom životu. Činjenica je da su turističke organizacije počele da rade ankete na ovim mikroblogovima da bi se dobio osećaj opšte sentimentalnosti za njihovu ponudu. Često ove organizacije studiraju korisnikove reakcije i odgovaraju korisnicima na mikroblogovima. IBM tehnologija pomoću modela za klasifikaciju tvitova u pozitivna, negativna i neutralna osećanja, može pomoći da se nastava osavremeni na nekim predmetima na smeru “turizam”.



Slika 45. Prikaz sentimentalne analize

U naprednoj verziji je moguće analizirati email adresu i pritom doći do određene statistike koja može biti od pomoći u poboljšanju kvaliteta nastave. Postoji dosta aplikacija koje takođe mogu da analiziraju email adresu, a koje su besplatne i luke za korišćenje. Za primer su uzete dve aplikacije luke za korišćenje – Immersion i Gmail Meter. Za vizualizaciju i statistiku je najbolje koristiti neku od aplikacija, a za analizu teksta IBM BigInsights.

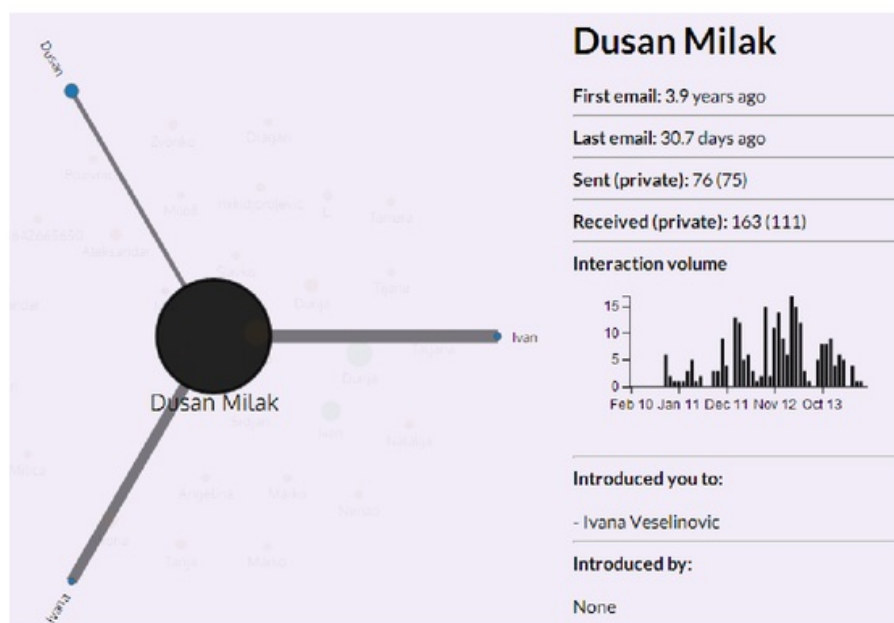
Aplikacija “Immersion” analizira podatke sa e-maila i brzo i jednostavno daje rezultate [71]. Pokazuje kome je poslato najviše email poruka, i od koga je koliko primljeno. Kada se pređe mišem preko krugova prikaže se broj mail-ova a za detaljnija objašnjenja dovoljno je kliknuti na njih. Takođe se može prikazati statistika za određenog korisnika i ako postoje veze sa drugim korisnicima.

Pomoću aplikacije Gmail Meter moguće je dobiti uvid u statistiku koja je vezana za dnevni, nedeljni i mesečni saobraćaj. Prikazani su dijagrami koji pokazuju kategoriju mejla i količinu teksta koji je primljen u odnosu na onaj koji je poslat. Dostupne su teme koje su najpopularnije kao i naslovi koji su najčešći. Prikazan je dijagram koji označava vreme kad je najviše email-a stiglo ili poslato, kao i grafikon koji pokazuje vrstu fajlova koji su pristizali [72].

Ove aplikacije imaju mogućnost da eksportuju mail-ove u tekstualni fajl koji može da se analizira preko IBM BigInsights-a, tačnije preko aplikacije „Word Count“ koja služi za prebrojavanje reči koje su se najviše koristile u mejlovima. Uz pomoć ovoga profesor može (posle završenog semestra) da analizira svoju email adresu i da ima uvid u to šta su ga studenti najčešće pitali, koja je tema bila najaktuelnija i koji su problemi bili najzastupljeniji. Sa ovim podacima profesor može prilagoditi gradivo za sledeću generaciju, i samim tim unaprediti svoja predavanja.



Slika 46. Vizuelni prikaz analize email-a



Slika 47 Detaljan prikaz korisnika od koga je primljeno najviše mail-ova

Još jedan od načina kako može da se koristi unapređena verzija IBM tehnologije je analiza studentskih tvitova, koji koriste određeni heštag u konverzaciji koja je vezana za predmet. Da bi se ova analiza uspešno primenila, potreban je dogovor sa grupom studenata (primer: studenti iz predmeta „informacione tehnologije u biznisu“) da se na Twitteru vode što češće konverzacije o predmetu u toku pripreme za kolokvijum ili ispit. Dogovor sa studentima bi bio da posle svakog

postavljenog pitanja koje je vezano za predmet, ili posle svake izjave, pomoći ili odgovora na nečije pitanje, tvit završi heštagom *#bpsitb* (*Beogradska poslovna škola - informacione tehnologije u biznisu*) radi lakšeg pristupa podacima. Nakon toga potrebno je izvršiti upit u Twitter aplikaciji koji izvlači sve tvitove koji u sebi imaju heštag *#bpsitb*. Upit bi izgledao ovako:

```
q=%23bpsitb&result_type=recent
```

Kao što je ranije opisano, Twitter na osnovu upita generiše kod koji se koristi da bi preko terminala izvukli tj. kreirali fajl sa ekstenzijom json iz kog se preko Eclipse kreiraju fajlovi za analizu. Pored fajla sa ekstenzijom del iz kog se prave tabele i analiziraju podaci u BigSheets-u, Eclipse pravi i tekstualni fajl koji služi za tekstualnu analizu i analizu brojanja reči. Kod u Eclipse izgleda ovako:

```
// TODO: Add JAQL content here
```

```
tweets = jaqlGet("file:///home/biadmin/bpsitb.json");
```

```
...
```

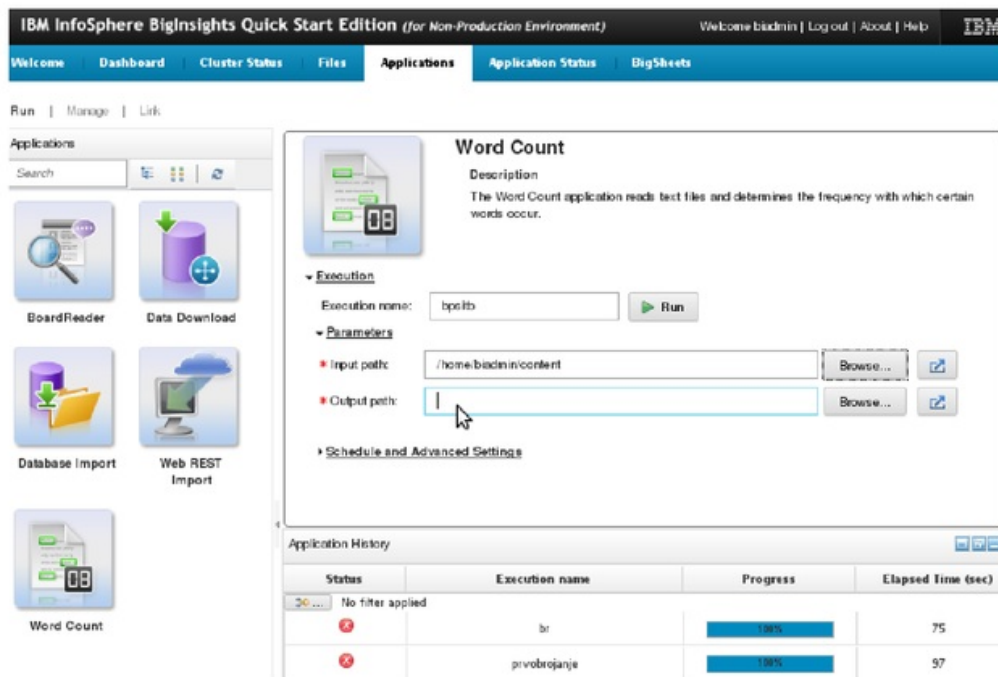
3

```
tweets_text = tweets -> transform $.text;
```

```
tweets_text -> write(seq("file:///home/biadmin/content/bpsitb.seq"));
```

Da bi se najbolje analizirali tvitovi koji su korišćeni, potrebno je prebrojati reči koje su najčešće korišćene, i na osnovu toga videti o čemu se najviše pisalo i šta može da predstavlja potencijalan problem studentima.

BigInsights uključuje dva test uzorka aplikacije popularnih u Hadoop - baziranim okruženjima: TeraGen - TeraSort i Word Count. TeraSort je MapReduce aplikacija koja radi paralelna sklapanja i sortiranja po ključevima u skupu podataka koji generiše TeraGen. To je polazna tačka koja kombinuje testiranje HDFS i MapReduce slojeva Hadoop klastera [73]. WordCount obrađuje kolekciju tekstualnih fajlova i vraća ukupan broj pojava svake reči koja se u tom fajlu nalazi. TeraGen - TeraSort generiše i sortira setove podataka čija se veličina meri terabajtima. Word Count je potrebno dodati preko taba Applications-Menage-Test-Word Count. Nakon otvaranja aplikacije potrebno je locirati tekstualni fajl, tj. folder u kojem se nalazi fajl koji treba da se analizira i izabрати folder gde će se upisati rezultati.



Slika 48. Pokretanje Word Count aplikacije

Nakon analize tvitova moguće je izabrati nekoliko vrsta vizualizacije, od kojih je najzanimljivija ta gde se reči koje su najčeće korišćene uvećavaju i suprotno. Rezultati koji se dobiju analizom se mogu iskoristiti tako što će se napraviti par tvitova koji će služiti kao odgovor studentima na postojeće probleme, i obratiti pažnja na to šta je studente najviše „kočilo“ dok su se spremali za kolokvijum, pa sledećoj generaciji skrenuti pažnju na eventualne probleme. Takođe treba više pažnje i vremena posvetiti tom delu gradiva.

Za primenu Twitter analize, na predmetima koji su vezani za informacione tehnologije, izabran je upit koji je vezan za android aplikacije. Profesor koji predaje android programiranje, neposredno pred predavanje može sa studentima da izvrši analizu Twitter podataka i da dobije informacije koje su u trendu a koje su vezene za tu oblast. Na osnovu tih informacija predavanja se mogu prilagoditi trendu. Potrebno je izvršiti upit u Twitter aplikaciji. Upiti mogu biti različiti a ovo su neki od primera (Tabela 6).

Upit:	Šta tvit sadrži:
watching now	Ovo je podrazumevana opcija, sadrži "watching" i "now"
"happy hour"	sadrži frazu "happy hour"
love OR hate	sadrži ili "love" ili "hate" (ili obe reči)
beer -root	sadrži beer ali ne i "root"
#haiku	sadrži heštag "haiku"
from:alexiskold	tvitovi koji su poslali od strane korisnika "alexiskold"
to:techcrunch	tvitovi koji su poslali ka korisniku "techcrunch"
superhero since:2010-12-27	tvitovi koji sadrže "superhero" i koji su poslali od datuma "2010-12-27" (godina-mesec-dan)
ftw unit:2010-12-27	tvitovi koji sadrže "ftw" ali koji su poslali pre navedenog datuma
movie -scary :)	sadrži "movie" ali ne i "scary", i izbacuje samo pozitivne tvitove
flight :(	sadrži "fight" i samo tvitove sa negativnim atributima
traffic ?	sadrži upitne rečenice sa "traffic"
hilarious filter:links	sadrži "hilarious" i linkove u tvitovima

Tabela 6. Različiti upiti za Twitter aplikaciju

Takođe je moguće tražiti upite za tvitove iz određenih lokacija, sa određenim jezikom koji je korišćen, moguće je tražiti tvitove koji su najpopularniji ili oni koji imaju više različitih upita, npr ako se postavi upit `"q=&geocode=-22.912214,-43.230182,1km&lang=pt&result_type=recent"` BigInsights će analizirati samo tvitove koji su skorije postavljeni, samo na portugalskom jeziku, sa lokacije - stadion Marakana, Rio, Brazil.

Konkretno za ovo istraživanje upit koji je kreiran izgleda ovako:

count=300&q=android+app&result_type=popular

Nakon toga se dobija fajl za analizu koji će sadržati tvitove koji sadrže "android app" i koji su popularni a ujedno i novi. Da bi dobili uvid u najaktuelnije android aplikacije, najbolje rešenje je sortirati tvitove po korisnicima koji su najuticajniji, tj. po korisnicima koji imaju najveći broj pratioca, zbog toga što se podrazumeva da ti korisnici uvek prvi objavljuju informacije o aktuelnim stvarima u svetu. Nakon tog sortiranja treba dodati i funkciju "URL" koja će iz tvitova u koloni 17 izvući samo linkove ka najaktuelnijim blogovima ili direktno ka aplikacijama koje su aktuelne.

Funkcije u BigSheets-u su unapred definisane formule koje vrše proračune pomoću specifičnih vrednosti ili argumente u određenom redosledu i proizvodi jednu povratnu vrednost. Sve BigSheets funkcije se nalaze u levom donjem uglu - "Function sheet". BigSheets je snabdeven sa nekoliko kompletnih funkcija koje su sasvim dovoljne za neku ne preterano zahtevnu analizu, ali može se instalirati i još dodatnih funkcija, kao i makro dodataka. Osnovne funkcije su podeljene u nekoliko grupa:

- Matematičke funkcije (Math) vrše zajedničke kalkulacije i operacije za rad sa brojevima.
- Funkcije teksta (Text function) vrše zajedničke operacije za manipulaciju i traženje teksta.
- Za prerađu dokumenata XML i HTML postoje funkcije: HTMLATTRVALUE, HTMLEXTRACTTAG, HTMLEXTRACTTAGS, HTMLREMOVEDMARKUP, HTMLTAGVALUE, XMLATTRVALUE, XMLEXTRACTTAG, XMLEXTRACTTAGS, XMLTAGVALUE.
- Vremenske funkcije (Date time) vrše zajedničke operacije za izračunavanje datuma i vremena. Ove funkcije mogu izračunati razlike između dva datuma, mogu izdvojiti godinu, datum ili nedelju iz DateTime objekta i mogu da konvertuju vremenske markere.

- Entitet funkcije (Entity functions) mogu izdvojiti određene entiete od izvornog tvita, npr. grad, država, ime, url itd.
- URL funkcije kupe informacije iz linkova u tvitovima. Mogu da naprave HTTP vezu sa navedenom URL adresom. Veze mogu da potraju neko vreme dok se ne završe, jer zavise od brzine mreže i odgovora iz URL. Neke od najpoznatijih url funkcija su: URLFRAGMENT, URLHOST, URLPATH, URLPORT, URLQUERY, URLPARAM, URLSCHEME, URLGET, URLGETHEADER, URLGETCONTENT.



Slika 49. Dodavanje funkcije “URL”

Posle dodavanja funkcije, pojavljuje se prozor gde je BigInsight izdvojio samo linkove iz tvitova koji su ih imali. Sledeći korak je ubacivanje kolone 17 koja sadrži kompletan tvit, da bi se imao bolji uvid u link koji je prikazan. Zatim kolona sa brojem pratioca, da bi se linkovi sortirali po korisnicima koji su najuticajniji, i na kraju kolona 15 koja prikazuje broj korisnika kojima se svideo sadžaj tvita. Posle sortiranja moguće je dobiti linkove koji su se najviše svideli korisnicima. Dobijena tabela u BigInsights je prikazana na slici 50.

IBM InfoSphere BigInsights Quick Start Edition (for Non-Production Environment)				
Welcome Dashboard Cluster Status Files Applications Application Status BigSheets				
Workbooks > View Results				
app(1)   Add chart + app > app(1) : build new workbook 				
Click run to update the data and the associated catalog table				
	url	header17	header12	header15
1	http://Lco/4QuimbZ8VA	本日 20 時より韓国祝日公式生にMSSPとこ	114950	398
2	http://Lco/EQm1YjMyu	本日 23 日 土曜日 20 時から韓国祝日の公式生に	126465	398
3	http://Lco/Y8kGwzLY	本日 23 日 土曜日 20 時から韓国祝日の公式生に	126465	398
4	http://Lco/GlwzyY7aey	本日 20 時より韓国祝日公式生にMSSPとこ	114950	398
5	http://Lco/ncQgovZV9	The #mZXFLUX app is available now for free c	2073899	225
6	http://Lco/6FCecJZLvj	The #mZXFLUX app is available now for free c	2073899	225
7	https://t.co/LH09UXZd	松田寛平君の黒子のバスケ担当編集か	30973	186
8	https://t.co/http://t.co/6FCecJZLvj	松田寛平君の黒子のバスケ担当編集か	30973	186
9	http://Lco/5QjeA2Bv	松田寛平君の黒子のバスケ担当編集か	30973	186
10	http://Lco/Er7WxWOrk5	Download our mobile apps now! iOS app at ht	87863	96
11	http://Lco/OCzjNPSQ	Download our mobile apps now! iOS app at ht	87863	96
12	http://Lco/1Bz7ydDq	Download our mobile apps now! iOS app at ht	87863	96
13	http://Lco/4SSfaGBW8E	The #mZXFLUX app is now available for free c	1372209	95
14	http://Lco/hSh2a3amle	The #mZXFLUX app is now available for free c	1372209	95
15	https://t.co/xleVsgdR1G	Para quem tem aparelho Android, o app do Bit	237208	68
16	http://Lco/3FGu8EgJXn	User beware: Researchers have 92% success	726689	42
17	http://Lco/E7Qat8SMom	User beware: Researchers have 92% success	726689	42
18	http://Lco/G0xaW4M5n	#Android weakness allows Hacking #Gmail and	151505	27
19	http://Lco/ncUumz9ncF	#Android weakness allows Hacking #Gmail and	151505	27
20	http://Lco/1GmmlEMM1K	Download FORGET IT! Android App Base knowledge	274931	9

Slika 50. Sortirani linkovi iz twitova koji su najpopularniji

IBM BigInsights daje mogućnost brzog otvaranja linka u tom prozoru kao brzi prikaz informacije sa te adrese, tako da korisnik ne mora da čeka otvaranje linka u novim prozorima. Ako se link čini zanimljivim, postoji opcija da se otvori u novom tabu. Posle otvorenih prvih 5 linkova, već može da se izvuče informacija koje su aplikacije najaktuelnije trenutno. Tvit koji je bio najviše favorizovan sadržao je link ka aplikaciji koja je izašla tog dana i zove se “Hacking Gmail App” i postavljena je na blogu “The hacker news”.

app(1): Open in new Window



Slika 51. Link koji se najviše svideo korisnicima u tom momentu

10.5. Izrada aplikacije

Aplikacija za analizu Twitter podataka je kreirana sa ciljem da olakša pretagu i pomogne profesorima da unaprede aktuelnost nastavnog sadržaja. Aplikaciju mogu da koriste i profesori koji nemaju puno iskustva u korišćenju računara. Dovoljno je da se startuje aplikacija, i popuniti par polja za pretragu. Rezultati će se prikazati u sledećem prozoru i biće dovoljni da profesor sazna šta je u trendu u toj oblasti ili da pronade tvitove najuticajnijih korisnika koji komentarišu određeni proizvod za koji je izvršena pretraga. Za realizaciju same aplikacije korišćen je PHP skript jezik.

Pristupanje Twitter API-ju pomoću PHP-a se vrši pomoću sledećeg koda:

Kod:

```
2
ini_set('display_errors', 1);

require_once('TwitterAPIExchange.php');

/** Set access tokens here - see: https://dev.twitter.com/apps/ **/
```

```

$settings = array(

    'oauth_access_token' => "xxxxxxxxxxxxxxxxxxxx",

    'oauth_access_token_secret' => "xxxxxxxxxxxxxxxxxxxx",

    'consumer_key' => "xxxxxxxxxxxxxxxxxxxx",

    'consumer_secret' => "xxxxxxxxxxxxxxxxxxxx"

);

/** URL for REST request, https://dev.Twitter.com/docs/api/1.1/ **/

$url = 'https://api.Twitter.com/1.1/blocks/create.json';

$requestMethod = 'POST';

/** POST fields required by the URL above. See relevant docs as above **/

$postfields = array(

    'screen_name' => 'usernameToBlock',

    'skip_status' => '1'

);

/** Perform a POST request and echo the response **/

$twitter = new TwitterAPIExchange($settings);

echo $twitter->buildOauth($url, $requestMethod)

    ->setPostfields($postfields)

    ->performRequest();

/** Perform a GET request and echo the response **/

/** Note: Set the GET field BEFORE calling buildOauth(). **/

$url = 'https://api.Twitter.com/1.1/followers/ids.json';

$getfield = '?screen_name=J7mbo';

$requestMethod = 'GET';

$twitter = new TwitterAPIExchange($settings);

```

```
echo $Twitter->setGetfield($getfield)
```

```
->buildOauth($url, $requestMethod)
```

```
->performRequest();
```

```
TwitterServiceProvider.php
1 <?php namespace Providers\Twitter;
2
3 use Guzzle\Plugin\Oauth\OauthPlugin;
4
5 class TwitterServiceProvider extends ServiceProvider
6 {
7     public function register() {
8         $this->app->bind('twitter', function()
9         {
10             $client = new Client('https://api.twitter.com/1.1');
11             $auth = new OauthPlugin([
12                 'consumer_key' => Config::get('twitter.consumer_key'),
13                 'consumer_secret' => Config::get('twitter.consumer_secret'),
14                 'token' => Config::get('twitter.token'),
15                 'token_secret' => Config::get('twitter.token_secret'),
16             ]);
17             $client->addSubscriber($auth);
18             return new TwitterAPI($client);
19         });
20     }
21 }
```

Slika 52. Registracija Twitter Service Provider funkcije

```
twitter.php
1 <?php
2 return array(
3     'consumer_key' => '90vvYtdNrTwCskphYCZRQzhKS',
4     'consumer_secret' => '8tuKz3vGE2spKvXyOQXOdW8GhyR5e0rFugIbKnTpgJHcdmXm2M',
5     'token' => '57613953-QN5mKeG56lx4FaIZuFlMmdQBgmLF2pTcURPXflxWv',
6     'token_secret' => 'PGUfAwXG13OjcOTYlk9otJOCajMgt9jnsHgAJ8M93h4Oz',
7     'oauth_token' => null,
8     'oauth_token_secret' => null,
9
10     'bearer_token' => null,
11
12     'callback_url' => url('twitter/callback'),
13     'fallback_url' => url('/')
14 );
```

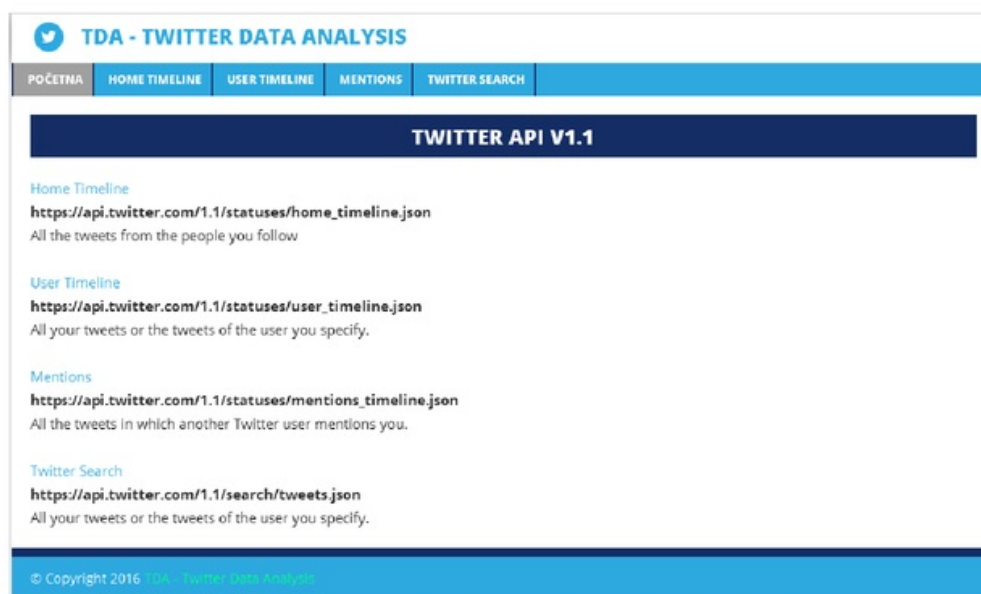
Slika 53. Upotreba tokena

```

1 <?php namespace Providers\Twitter;
2
3 use Guzzle\Service\Client;
4
5 class TwitterAPI {
6
7     protected $client_value;
8
9     public function __construct(Client $client_value)
10     {
11         $this->client=$client_value;
12     }
13
14     public function search($query)
15     {
16         $response = $this->$client_value->get("search/tweets.json?q=$query")->send();
17
18         return $response->json();
19     }
20 }
21

```

Slika 54. Prosledivanje parametara \$query



Slika 55. Grafički interfejs za prikupljanje podataka

Nakon startovanja aplikacije profesor treba da popuni par polja iz kojih će se napraviti upit koji će se dalje proslediti Twitter aplikaciji koja će na osnovu njega generisati kod koji je potreban za kreiranje fajla pogodnog za analizu u IBM BigInsights tehnologiji.

U polje za pretragu je potrebno uneti pojam ili frazu za koju se vrši analiza. Korisnik bira način na koji želi da mu se prikažu podaci, u mogućnosti je da izabere neki od ponuđenih modula za prikaz analize podataka. Moguće je izabrati prikaz tvitova iz određenog dela sveta. Potom treba čekirati jezik za koji se vrši pretraga, tj jezik na kojem je napisan tvit. Zatim zemlju porekla, i vrsta rezultata (analiza tvitova koji imaju samo pozitivan ili samo negativan karakter). Posle toga je polje za broj tvitova i prikaz tvitova do određenog perioda. Nijedno od ovih polja nije obavezno za popunjavanje, a moguće je i čekirati više opcija za pretragu. Na slici 56. je prikazan izgled aplikacije nakon popunjenih zahteva za analizu podataka.

The screenshot displays the 'TDA - TWITTER DATA ANALYSIS' web application. At the top, there is a navigation bar with tabs: 'POČETNA', 'HOME TIMELINE', 'USER TIMELINE', 'MENTIONS', and 'TWITTER SEARCH'. The 'TWITTER SEARCH' tab is currently selected. Below the navigation bar is a dark blue header with the text 'TWITTER SEARCH'. The main form area contains the following elements: a text input field with the URL 'https://api.twitter.com/1.1/search/tweets.json' and a hint 'All your tweets or the tweets of the user you specify.'; a label 'Pojam za pretragu q:' followed by a text input field containing 'android'; a label 'Geo kod:' followed by an empty text input field; a label 'Jezik:' followed by an empty text input field; a label 'Zemlja:' followed by an empty text input field; a label 'Vrsta Rezultata:' followed by a dropdown menu showing 'mixed'; a label 'Broj Tvitova:' followed by a dropdown menu showing '1000'; and a label 'Prikaz Tvitova do određenog perioda (YYYY-MM-DD):' followed by an empty text input field. At the bottom of the form are two large blue buttons: 'TRAŽI' and 'SAČUVAJ FAJL'.

Slika 56. Izgled aplikacije nakon popunjenih zahteva za analizu podataka

Rezultati koje aplikacija prikazuje posle upita koji je naveden, prikazani su na slici 57.

The screenshot shows the TDA - Twitter Data Analysis web application. At the top, there's a navigation bar with tabs: POČETNA, HOME TIMELINE, USER TIMELINE, MENTIONS, and TWITTER SEARCH. The 'TWITTER SEARCH' tab is active. Below it, a search bar contains the text 'android'. The results are displayed in a table with columns: url, header17, header12, and header15. The table shows 20 rows of data, including various URLs and text snippets related to Android. At the bottom, there's a footer with the copyright notice: © Copyright 2016 TDA - Twitter Data Analysis.

	url	header17	header12	header15
1	http://t.co/4QaimbZ8VIA	本日20時より韓国民舞公式生にMISPPとこ	114990	308
2	http://t.co/8QnYVMA4	本日20時より韓国民舞の公式生に	136405	308
3	http://t.co/Y8IGe9y4	http://t.co/4QaimbZ8VIA 本日20時より韓国民舞の公式生に	136405	308
4	http://t.co/GhezyYaqy	本日20時より韓国民舞公式生にMISPPとこ	114990	308
5	http://t.co/wCQazV9	The #m2XFLUX app is available now for free c 2073869	226	
6	http://t.co/8FCaoJZu1	The #m2XFLUX app is available now for free c 2073869	226	
7	https://t.co/94LH0GZd	松田寛平氏の真子の「スグ総合編集カ」	30673	196
8	https://t.co/Ha1Ju0dp	松田寛平氏の真子の「スグ総合編集カ」	30673	196
9	http://t.co/5Cjw428v	松田寛平氏の真子の「スグ総合編集カ」	30673	196
10	http://t.co/6r7W904S	Download our mobile apps now! iOS app at the	87963	96
11	http://t.co/0CLyNP0Q	Download our mobile apps now! iOS app at the	87963	96
12	http://t.co/98d7yd0q	Download our mobile apps now! iOS app at the	87963	96
13	http://t.co/458uQ8V8E	The #m2XFLUX app is now available for free c 1372209	95	
14	http://t.co/9G2d8w4e	The #m2XFLUX app is now available for free c 1372209	95	
15	http://t.co/v6v8g8R1G	Plus quam ten apudto Android, o app do Bit 257208	68	
16	http://t.co/3FQd8gXk	User browser: Researchers have 92% success 726989	42	
17	http://t.co/ETQn85H4m	User browser: Researchers have 92% success 726989	42	
18	http://t.co/0DaKa4M5n	#Android weakness allows Hacking #Gmail and 151505	27	
19	http://t.co/wcUuz9ncF	#Android weakness allows Hacking #Gmail and 151505	27	
20	http://t.co/LGggl89tK	download GRATIS Android App Beaswario c 27530	3	

Preview of 50 rows from 100 [Previous](#) [Next](#)

© Copyright 2016 TDA - Twitter Data Analysis

Slika 57. Rezultati nakon upita

10.6. Upitnik

Sa osnovnim ciljem da se ispita učestalost korišćenja informaciono-komunikacionih tehnologija u edukaciji u Beogradskoj poslovnoj školi na smeru „finansije i računovodstvo“ i „marketing i trgovina“ i kreiran je prvi upitnik koji su popunjavali studenti sa tih smerova u toku studija druge godine četvrtog semestra. Takođe je kreiran i drugi upitnik koji su popunjavali studenti koji koriste informaciono-komunikacione tehnologije u obrazovne svrhe. Drugi upitnik je kreiran sa ciljem da se stekne uvid u to šta studenti sa pomenutih smerova misle o značaju informacionih tehnologija u poboljšanju edukativnog procesa u Beogradskoj poslovnoj školi.

Sprovedeno je istraživanje sa uzorkom od 100 studenata koji su popunjavali upitnik. On je kreiran sa ciljem da se izvrši ispitivanje studenata o njihovoj percepciji, efektima i uticajima informacionih tehnologija na kvalitet edukativnog procesa. Podaci su obrađeni i analizirani i rezultati ispitivanja vrlo jasno ukazuju na pozitivan stav ispitanika.

Prvi upitnik sadrži devet pitanja zatvorenog tipa i uglavnom se odnose na utvrđivanje koliko su ispitanici upoznati sa savremenim edukacionim modelima. Izgled prvog upitnika:

UPITNIK O KORIŠĆENJU INFORMACIONIH TEHNOLOGIJA U EDUKACIJI U BEOGRADSKOJ POSLOVNOJ ŠKOLI NA SMERU "FINANSIJE I RAČUNOVODSTVO" I "MARKETING I TRGOVINA"

1

Poštovani studenti i studentkinje, pred Vama se nalazi upitnik koji je kreiran u cilju istraživanja koliko studenti koriste informacione tehnologije u obrazovne svrhe.

1

Molim Vas da pročitate svako pitanje i da svojim iskrenim odgovorima doprinesete dobijanju što objektivnijih podataka koji će biti upotrebljeni samo za naučnu analizu. Upitnik je potpuno anoniman i rezultati će se koristiti samo u istraživačke svrhe.

Unapred Vam se zahvaljujem na saradnji i spremnosti da svojim učešćem doprinesete kvalitetnijoj realizaciji ovoga istraživanja.

1. Da li ste nekada koristili informaciono-komunikacione tehnologije u obrazovne svrhe?
 - Da
 - Ne
2. Da li koristite e-mail adresu u svojstvu komunikacije sa profesorom ili asistentom?
 - Da
 - Ponekad (retko)
 - Ne
3. Da li koristite Skype za komunikaciju sa profesorom ili asistentom ukoliko niste u mogućnosti da prisustvujete konsultacijama?
 - Da
 - Ponekad (retko)
 - Ne

4. Da li koristite projektor ili kompjuter kao pomoć pri odbrani seminarskog rada?

- Da

- Ponekad (retko)

- Ne

5. Da li koristite video materijal sa održanih predavanja ukoliko ste ih propustili?

- Da

- Ponekad (retko)

- Ne

6. Da li ste nekad odslušali besplatan kurs na internetu?

- Da

- Ponekad (retko)

- Ne

7. Da li koristite pametan telefon kao pomoć u edukaciji?

- Da

- Ponekad (retko)

- Ne

8. Da li ste korisnik nekog virtualna okruženja?

- Da

- Ne

9. Da li ste korisnik neke društvene mreže?

- Da

- Ne

Drugi upitnik su popunjavali studenti koji koriste informaciono-komunikacione tehnologije u obrazovne svrhe. Od njih 100 koji su popunjavali prvi upitnik, njih 91 je popunilo i drugi što govori o tome da ogroman procenat studenata koristi barem neke informacione tehnologije koje im olakšavaju školovanje. Za potrebe istraživanja efekata koje imaju informaciono-komunikacione tehnologije na visoko obrazovanje, konkretno efekta koje imaju na poboljšanje kvaliteta nastave u Beogradskoj poslovnoj školi, kreirano je osam pitanja. Kao odgovor na dva ponuđena pitanja konstruisana je četvorostepena skala. U opisu pitanja ispitanicima je objašnjeno da se skalom nastoji ispitati njihovo mišljenje o uticaju informacionih tehnologija u poboljšanju edukativnog procesa u Beogradskoj poslovnoj školi na smeru „finansije i računovodstvo“ i „marketing i trgovina“. Zadatak ispitanika bio je da navedu u kojoj meri navedene stavke mogu olakšati proces školovanja (1 – uopšte ne olakšava; 4 – olakšava i te kako;). Ostalih šest pitanja je bilo zatvorenog tipa a kreirani su sa istim ciljem. Izgled drugog upitnika:

UPITNIK O ZNAČAJU INFORMACIONIH TEHNOLOGIJA U POBOLJŠANJU EDUKATIVNOG PROCESA U BEOGRADSKOJ POSLOVNOJ ŠKOLI NA SMERU "FINANSIJE I RAČUNOVODSTVO" I "MARKETING I TRGOVINA"

Poštovani studenti i studentkinje, pred Vama se nalazi upitnik koji je kreiran u cilju istraživanja da se stekne uvid u mišljenje studenata o tome koliko informacione tehnologije utiču na kvalitet nastavnog procesa.

Molim Vas da pročitate svako pitanje i da svojim iskrenim odgovorima doprinesete dobijanju što objektivnijih podataka koji će biti upotrebljeni samo za naučnu analizu. Upitnik je potpuno anoniman i rezultati će se koristiti samo u istraživačke svrhe.

Unapred Vam se zahvaljujem na saradnji i spremnosti da svojim učešćem doprinesete kvalitetnijoj realizaciji ovoga istraživanja.

1. Da li Vam informacione tehnologije olakšavaju proces školovanja?

- da
- mogu i bez njih
- ne

2. Označite u kojoj mjeri ¹Vam savremeni načini komunikacije sa profesorima olakšavaju proces školovanja? (Označite ocenu koja najbolje odražava vaš stepen slaganja sa navedenom tvrdnjom: 1 – uopšte ne olakšava; 2 – možda olakšava, ali ne koristim; 3 – olakšava; 4 – olakšava i te kako;)

	1	2	3	4
Telefonski poziv ili SMS	☐	☐	☐	☒
E-mail	☐	☐	☐	☒
Viber	☐	☐	☐	☒
Skype	☐	☐	☐	☒
Facebook messenger	☐	☐	☐	☒

3. Da li online praćenje nastave preko IP kamere može da poboljša kvalitet obrazovanja?

- Ne može, tradicionalni pristupi su mnogo efikasniji
- Verovatno može, ali ja nemam mogućnosti da ih pratim
- Može, rado pratim online predavanja

4. Da li besplatni online kursevi mogu da budu od pomoći pri učenju?

- Ne
- Možda, ali ne praktikujem takvu vrstu učenja
- Da, i te kako

5. Da li pametni mobilni telefoni mogu da budu od pomoći pri učenju?

- Ne
- Možda, ali ne praktikujem takvu vrstu učenja
- Da, i te kako

6. Koliko korišćenje virtualnih okruženja ili savremenih android aplikacija mogu da pomognu pri učenju?

- Nimalo
- Možda mogu, ali ja ih ne praktikujem
- Pomažu i te kako

7. Ocenite u kojoj meri društvene mreže mogu da olakšaju proces školovanja? (Označite ¹ocenu koja najbolje odražava vaš stepen slaganja sa navedenom tvrdnjom: 1 – uopšte ne olakšava; 2 – možda olakšava, ali ne koristim; 3 – olakšava; 4 – olakšava i te kako;)

	1	2	3	4
Facebook	☐	☐	☐	☐
Twitter	☐	☐	☐	☐
Youtube	☐	☐	☐	☐
Instagram	☐	☐	☐	☐
Google +	☐	☐	☐	☐

8. Da li aplikacija TDAedu može da osavremeni nastavni proces?

- Ne
- Ne znam, nisam upućen/a
- Da

Rezultati prvog upitnika:

QUESTIONS

RESPONSES

100

100 responses



SUMMARY

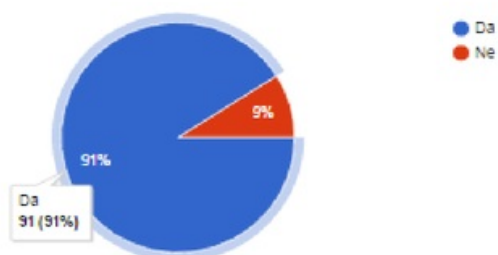
INDIVIDUAL

Accepting responses



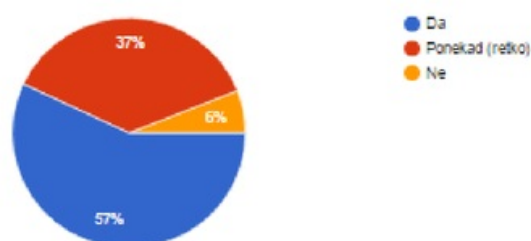
Da li ste nekada koristili informaciono-komunikacione tehnologije u obrazovne svrhe?

(100 responses)



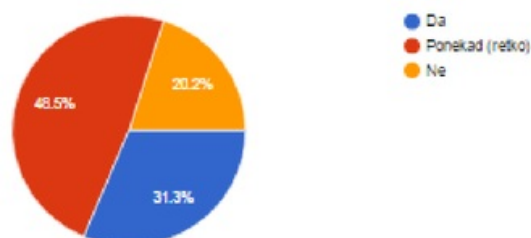
Da li koristite e-mail adresu u svojstvu komunikacije sa profesorom ili asistentom?

(100 responses)



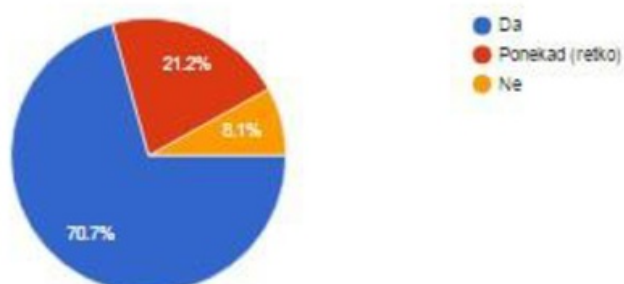
Da li koristite Skype za komunikaciju sa profesorom ili asistentom ukoliko niste u mogućnosti da prisustvujete konsultacijama?

(99 responses)



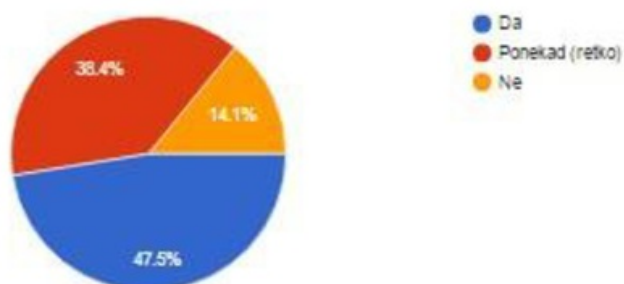
Da li koristite projektor ili kompjuter kao pomoć pri odbrani seminarskog rada?

(99 responses)

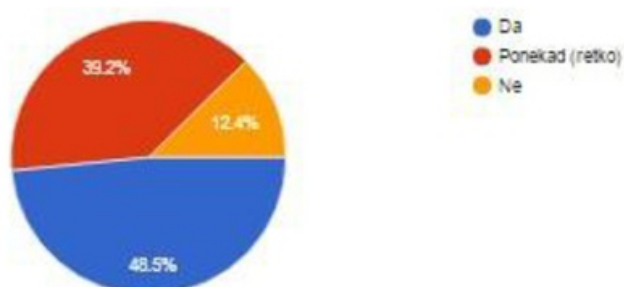


Da li koristite video materijal sa održanih predavanja ukoliko ste ih propustili?

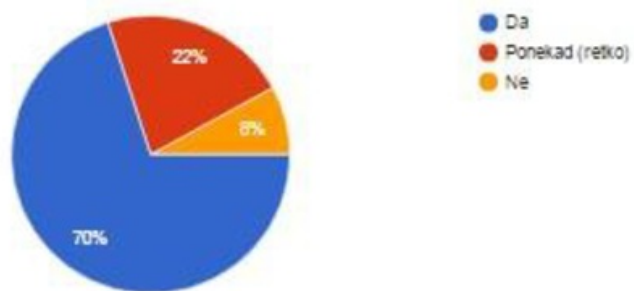
(99 responses)



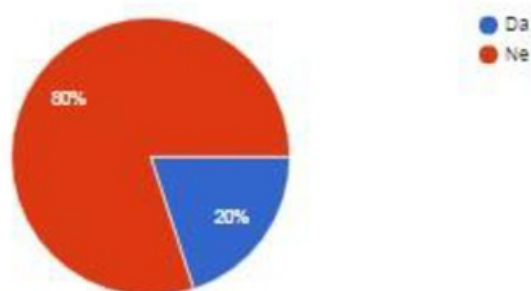
Da li ste nekad odslušali besplatan kurs na internetu? (97 responses)



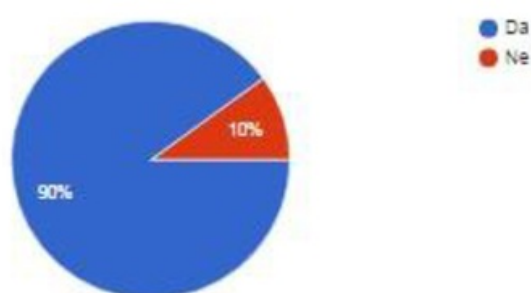
Da li koristite pametan telefon kao pomoć u edukaciji? (100 responses)



Da li ste korisnik nekog virtualna okruženja? (100 responses)



Da li ste korisnik neke društvene mreže? (100 responses)



Slika 58. Statistički prikaz rezultata prvog upitnika

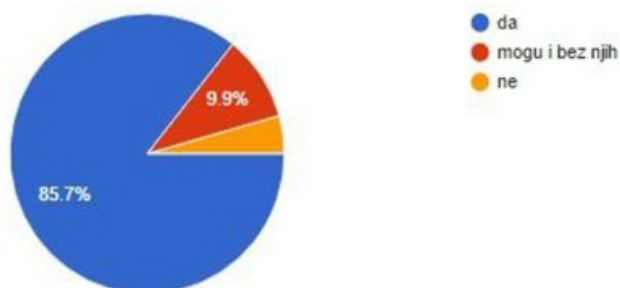
Dve grupe studenata su popunjavale ovaj upitnik i po tome je on specifičan i značajan. Prva polovina ispitanika su studenti koji su trenutno upisani na studijski program „finansije i računovodstvo“ u Beogradskoj poslovnoj školi, a druga polovina su studenti druge godine sa programa „marketing i trgovina“. Ukupan broj studenata koji je učestvovao u ovom istraživanju je 100. Svi su popunjavali prvi upitnik, a drugi je popunjavalo 91 ispitanika. Upitnik je realizovan u periodu od oktobra do novembra 2015. godine.

Analizom prvog upitnika moguće je zaključiti da 91% ispitanih studenata u nekoj meri koristi informacione tehnologije u obrazovne svrhe. Zanimljiv je podatak da čak 90% ispitanika koristi neku društvenu mrežu. Većina njih koristi mobilne telefone, komjutere i projektore kao pomoć u edukaciji. Najmanji procenat je vezan za korišćenje virtualnih svetova u edukativne svrhe, a tek 50% je nekad koristilo besplatne masovne onlajn kurseve. Kao moguće objašnjenje ovakvog stava studenata, nameće se podatak da je još uvek u zvaničnim statističkim podacima Coursera-e i EDX-a mali procenat polaznika iz Srbije. 31% koristi a 48% koristi ponekad Skype u nastavi.

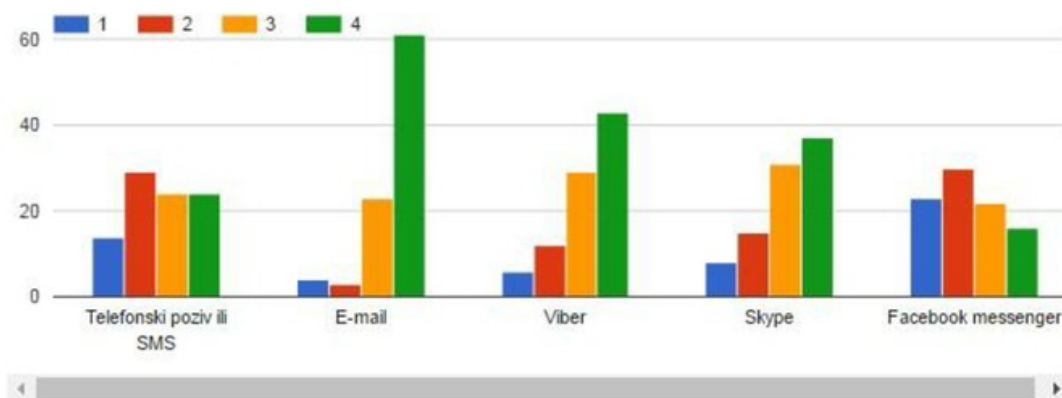
Rezultati drugog upitnika:



Da li Vam informacione tehnologije olakšavaju proces školovanja? (91 responses)

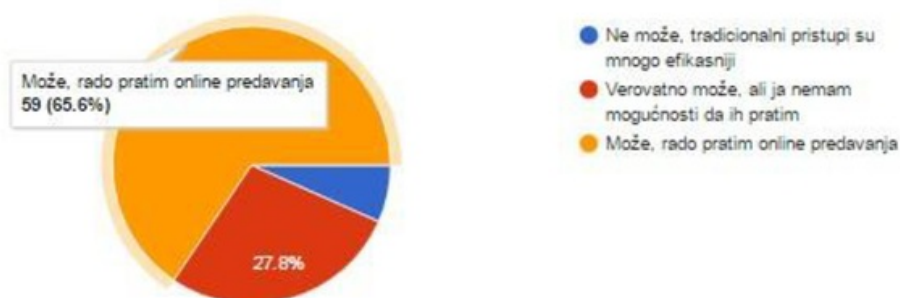


Označite u kojoj meri Vam savremeni načini komunikacije sa profesorima olakšavaju proces školovanja? (Označite ocenu koja najbolje odražava vaš stepen slaganja sa navedenom tvrdnjom: 1 – uopšte ne olakšava; 2 – možda olakšava, ali ne koristim; 3 – olakšava; 4 – olakšava i te kako;)



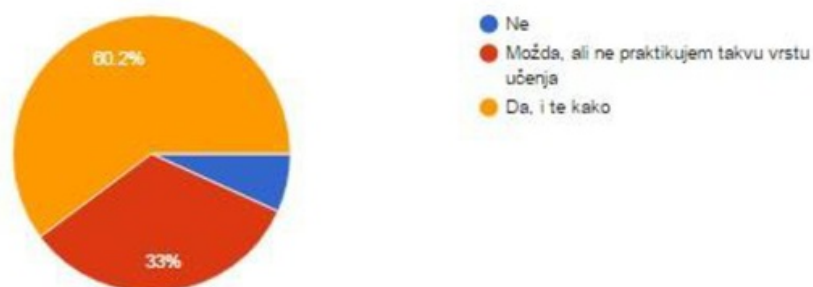
Da li online praćenje nastave preko IP kamere može da poboljša kvalitet obrazovanja?

(90 responses)



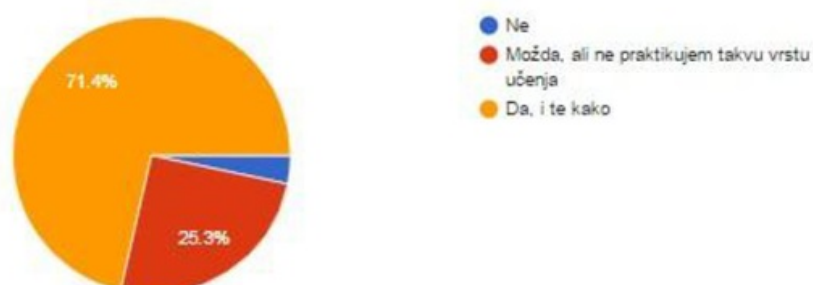
Da li besplatni online kursevi mogu da budu od pomoći pri učenju?

(88 responses)



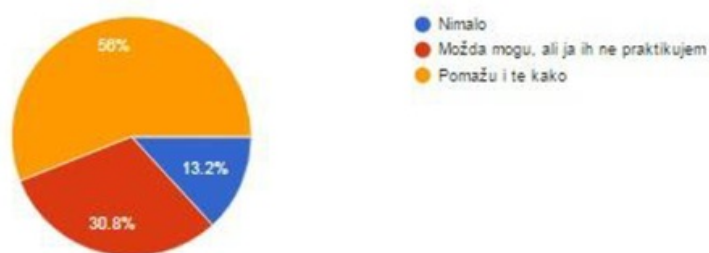
Da li pametni mobilni telefoni mogu da budu od pomoći pri učenju?

(91 responses)

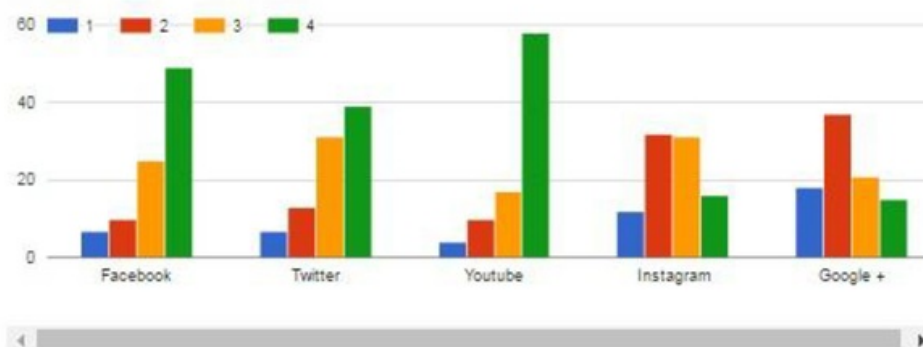


Koliko korišćenje virtualnih okruženja ili savremenih android aplikacija mogu da pomognu pri učenju?

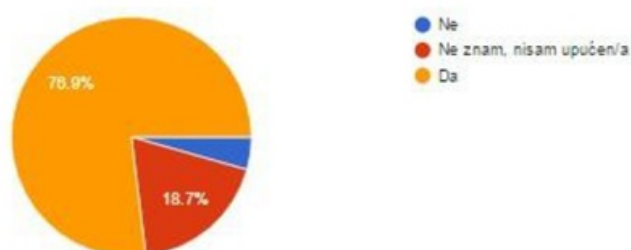
(91 responses)



Ocenite u kojoj meri društvene mreže mogu da olakšaju proces školovanja? (Označite ocenu koja najbolje odražava vaš stepen slaganja sa navedenom tvrdnjom: 1 – uopšte ne olakšava; 2 – možda olakšava, ali ne koristim; 3 – olakšava; 4 – olakšava i te kako;)



Da li aplikacija TDAedu može da osavremeni nastavni proces? (91 responses)



Slika 59. Statistički prikaz rezultata drugog upitnika

U drugom upitniku su pitanja bile više okrenuta ka tome ¹ da se stekne uvid u mišljenje studenata o uticaju informaciono-komunikacionih tehnologija na edukativni proces. Rezultati prvog pitanja pokazuju da je samo 4,4% ispitanih odgovorilo negativno na pitanje da li im informacione tehnologije olakšavaju proces školovanja. Što se tiče savremenih načina komunikacije koje olakšavaju proces školovanja, studenti su uglavnom imali pozitivan stav o *E-mail-u*, za nje su manje o *Viber-u* i *Skype-u*, a Facebook poruke, sms i telefonski pozivi su dobili lošiju ocenu. ¹ Kao moguće objašnjenje ovakvog stava studenata, nameće se činjenica da profesori i asistenti retko daju svoj broj telefona ili Facebook, uglavnom zbog želje da zaštite svoju privatnost. U narednim odgovorima studenti su iznosili mišljenja da li video materijali i besplatni onlajn kursevi mogu da poboljšaju kvalitet obrazovanja. Više od 60% ispitanika je uglavnom odgovorilo pozitivno. 71,4% ispitanika se slaže da pametni mobilni telefoni i te kako pomažu pri učenju, a najmanji procenat misli da virtualno okreženje može doprineti poboljšanju nastavnog procesa, uglavnom većina nije upoznata sa tim načinom učenja. S obzirom da je u prethodnom upitniku 90% ispitanika odgovorilo pozitivno na pitanje da li koriste društvene mreže, interesantno je kako su ocenili u kojoj meri one mogu da olakšaju proces školovanja. Najpozitivniji stav studenti su imali prema *Youtube-u*, Facebook-u i Twitter-u, dok Instagram i Google+ su mreže koje studenti najmanje upotrebljavaju i samim tim su najviše davali ocenu 2 (možda, ali ne koristim). Negativne ocene su procentualno zanemarljive, jer većina njih uopšteno misli da društvene mreže i te kako pomažu i imaju uticaj na sam kvalitet obrazovanja u visokoškolskim ustanovama. Zbog nekih ranijih istraživanja i testiranja aplikacije *TDAedu*, studenti su donekle bili upoznati sa radom te aplikacije kao i njenom uticaju na kvalitet predavanja. S obzirom na tu činjenicu, iskorišćena je prilika da se u upitniku 2 nađe pitanje koje može da da uvid u to šta studenti zapravo misle o pomenutoj aplikaciji i da li veliki podaci preuzeti sa tvitera mogu obradom da osavremene nastavni proces. 76,9% ispitanih studenata je odgovorilo pozitivno, 4,4% negativno a 18,7% nije upoznato sa radom aplikacije.

Gledajući uopšteno, rezultati ovih upitnika pokazuju da studenti Beogradske poslovne škole veoma često koriste neke od informaciono-komunikacionih tehnologija kako bi olakšali učenje i unapredili i osavremenili svoje edukativne veštine. Ispitanici imaju vrlo pozitivan stav što se tiče društvenih mreža u edukaciji i pozitivno reaguju na nove tehnologije koje se sve češće koriste u komunikaciji između nastavnog osoblja i studenata.

11. PREGLED DOPRINOSA DISERTACIJE

Rezultati istraživanja u ovom doktorskom radu su objavljeni u tri časopisa međunarodnog značaja (dva rada su sa SCI liste klasifikovana kao M23 od strane Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije) i saopšteni na jednoj međunarodnoj konferenciji. Na taj način je i zvanično izvršena verifikacija rezultata istraživanja od strane domaće i strane naučne javnosti.

1. Milovanović E., Popović M., Stanković R., Milovanović I., Remark on ordinary and Randić energy of graph, Journal of Mathematical Inequalities, dx.doi.org/10.7153/jmi-10-55, pp. 687-692, 2015 (M23)
2. Popović M., Stanković R., Milovanović I., Milovanović E., Some Remarks on Normalized Laplacian Spreads of Graphs, Revista Tecnica De La Facultad De Ingenieria Universidad Del Zulia; ISSN: 0254-0770, Facultad de Ingeniería, Universidad del Zulia, Vol. 38, N° 2, pp. 13-17, 2015 (M23)
3. Milovanović I., Milovanović E., Popović M., Stanković R., Remark on the Laplacian-energy-like and Laplacian incidence energy invariants of graphs, Creative Mathematics and Informatics, Ms. Ref. No.: CMI-15-01303R1, ISSN 1843-441X, No. 2, pp. 183-187, 2015
4. M. Popović, M. Milosavljević, "Twitter Data Analytics in Education Using IBM Infosphere Biginsights," in Sinteza 2016 - International Scientific Conference on ICT and E-Business Related Research, Belgrade, Singidunum University, Serbia, pp. 74-80. doi:10.15308/Sinteza-2016-74-80, 2016

Rezultati istraživanja u ovom radu se mogu opisati kroz sledeće naučne doprinose:

- Sistematizacija i klasifikacija postojećih rešenja savremenih edukacionih sistema;
- Analiza loših osobina postojećih modela i sistema;
- Analiza savremenih tehnologija za realizaciju softverskog rešenja;
- Definicije granica određenih parametara grafa koristeći spektralnu teoriju grafova i primena dobijenih rezultata u analizi Twitter društvene mreže;
- Predlog novog edukativnog okruženja koje omogućuje osavremenjivanje nastavnog sadržaja uz pomoć Twitter društvene mreže i podataka dobijenih i analiziranih od strane Big data tehnologije;

- Realizacija aplikacije edukacionog sistema TDAedu i integracija na različitim smerovima u Beogradskoj poslovnoj školi;
- Analiza primene predloženog rešenja;
- Standardizacija procesa projektovanja koncepta učenja uz pomoć društvenih mreža;
- Predlog budućih pravaca rada na osnovu prethodnog istraživanja.

12. ZAKLJUČAK

U radu je kroz primere pokazano da se implementacijom savremenih internet tehnologija u nastavni proces može doprineti boljem usvajanju gradiva. Sa ciljem da se poveća efikasnost učenja i podučavanja obrazovni proces se mora obogatiti atraktivnijim sadržajem koji bi studente uključio u proces obrazovanja kao aktivne učesnike. Personalizovani pristup u predavanju baziran na podacima dobijenih putem društvenih mreža daje bolje rezultate. Razmena podataka i iskustva sa socijalnih mreža pomaže u rešavanju konkretnih problema u edukaciji. Efikasnost, konkurentnost i efektivnost visokoškolskih ustanova zavisiće uglavnom od njihovih mogućnosti da se prilagode navikama i zahtevima studenata i da sprovedu nove metode koje bi unapredile performanse učenja. Implementirana tehnika za analizu društvene mreže Twitter doprinosi opštem poboljšanju rezultata studenata kao i pozitivnijem odnosu sa profesorima. Aplikacija za analizu Twitter podataka TDAedu se uspešno primenjuje u Beogradskoj poslovnoj školi za realizaciju nastavnog procesa. Time se potvrđuje osnovna hipoteza ovog istraživanja da novi način za unapređenje performansi obrazovnih sistema može biti implementacija informacija dobijenih analizom podataka sa društvene mreže Twitter, pomoću IBM BigInsights Big data tehnologije. Nakon što je primenjen i analiziran upitnik o stavu studenata prema uticaju informaciono-komunikacionih tehnologija na poboljšanje kvaliteta nastavnog procesa, donešen je zaključak da su ispitani studenti pozitivno reagovali na sve nove tehnologije koje su opisane u radu. Kao dobar primer se može navesti to da je 71,4% ispitanika odgovorilo pozitivno na pitanje da li im pametni mobilni telefoni pomažu pri učenju. Dobar deo ispitanika (76,9%) je odgovorio pozitivno i na pitanje koje se odnosilo na primenu aplikacije TDAedu u osavremenjivanju edukacije. Aplikacija je ranije prikazana i testirana pred većinom ispitanika tako da su studenti uglavnom bili upoznati sa radom te aplikacije kao i njenim uticajem na kvalitet predavanja.

Rezultati ovog istraživanja otvaraju mogućnosti za dalji rad u ovoj oblasti kroz nadogradnju i proširenje predloženih rešenja, kreiranje i razvoj kurseva elektronskog učenja na osnovu podataka o ponašanju i aktivnostima studenata kako na fakultetu, tako i na društvenim mrežama. Ti podaci bi se u novom softverskom sistemu obrađivali i primenjivali u realnom vremenu. Cilj budućih istraživanja bi bio da se razvije aplikacija za kreiranje adaptivnih kurseva koji bi bili prilagođeni profilima studenata i profesora. Predloženo rešenje buduće aplikacije bi trebalo da prati razvoj Twitter servisa da bi se podigao kvalitet upita i ponuđenih opcija. Posebna pažnja biće poklonjena razvoju modula softverskog sistema za interaktivnost i vizualizaciju.

13. LITERATURA

- [1] Stela Filipi Matutinovi, Naučne informacije u Srbiji - Protok, dostupnost, vrednovanje, 3. izmenjeno i dopunjeno izdanje, Beograd, 2014
- [2] M. Popović, M. Milosavljević, "Twitter Data Analytics in Education Using IBM Infosphere Biginsights," in Sinteza 2016 - International Scientific Conference on ICT and E-Business Related Research, Belgrade, Singidunum University, Serbia, pp. 74-80. doi:10.15308/Sinteza-2016-74-80, 2016
- [3] Takahiro Yoshida, Seiichiro Hangai, Development and Evaluation of a Keyword-Accessible Lecture Video Player and Lecture Video Contents, Strategic applications of distance learning technologies, Information Science Reference, New York, 2009
- [4] Rick Krueger, There are so many more than three barriers, Encyclopedia of Distance Learning, Second Edition, Information, Science Reference, New York, USA, 2009
- [5] Zoran Vučetić, Mogućnosti primene mobilnih telefona u obrazovanju, Tehnika i informatika u obrazovanju, Konferencija Čačak, 2008
- [6] Siniša G. Minić, Tehnologija, informatika i obrazovanje za društvo učenja i znanja, 6. Međunarodni Simpozijum, Tehnički fakultet Čačak, 2011
- [7] <http://www.raf.edu.rs/citaliste/operativni-sistemi/3657-istorija-mobilnih-os>

(Link posećen: 12.09.2016.)

[8] [http://web2tools-technologies.wikispaces.com/M-u%C4%8Denje%20\(mobile%20learning\)](http://web2tools-technologies.wikispaces.com/M-u%C4%8Denje%20(mobile%20learning))

(Link posećen: 08.07.2016.)

[9] Daniel C. Cliburn, Jeffrey L. Gross, Second Life as a Medium for Lecturing in College Courses, IEEE Computer Society, 978-0-7695-3450-3/09.

[10] http://wiki.secondlife.com/wiki/Second_Life_Education_Directory

(Link posećen 22.04.2014.)

[11] <http://openwonderland.org/> (Link posećen 28.07.2014.)

[12] Fabio Ricardo dos Santos, Christian Guetl, Philip H. Bailey, V. Judson Harward, Dynamic Virtual Environment for Multiple Physics Experiments in Higher Education, 978-1-4244-6571/2/10 IEEE Educon 2010

[13] <http://www.infotrend.hr/clanak/2013/11/big-data---izazovi-implementacije,78,1029.html>

(Link posećen: 20.04.2014.)

[14] Živko Krstić, Big data, jeste li spremni, Udruženje studenata informatičkog menadžmenta, Beograd, 2012

[15] Hanghang Tong, Optimal Connectivity on Big Graphs: Measures, Algorithms and Applications, Data lab, Arisona State University, 2015

[16] http://www.novell.com/products/file-reporter/features/accurate_assessment.html

(Link posećen: 18.03.2015.)

[17] <http://hackeducation.com/2012/12/09/top-ed-tech-trends-of-2012-education-data-and-learning-analytics/> (Link posećen: 22.06.2014.)

[18] Xiongpai QIN, Huiju WANG, Furong LI, Baoyao ZHOU, Cuiping LI, Hong CHEN, Xuan ZHOU, Xiaoyong DU, Yu CAO, Shan WANG, Beyond Simple Integration of RDBMS and MapReduce – Paving the Way toward a Unified System for Big Data Analytics: Vision and Progress, - 2012 IEEE - DOI 10.1109/CGC.2012.39

[19] <http://www.wired.com/wiredscience/2013/10/big-data-science/all/> (Link posećen: 15.06.2014.)

[20] <http://pcpress.rs/ibm-big-data/> (Link posećen: 01.07.2014.)

- [21] <http://www.mindcet.org/en/report/big-data-education-2> (Link posećen: 01.07.2014.)
- [22] <http://www.kairos3d.it/solutions/big-data/> (Link posećen: 01.11.2014.)
- [23] <http://polymaps.org/> (Link posećen: 01.11.2014.)
- [24] <http://www.q-researchsoftware.com/> (Link posećen: 01.11.2014.)
- [25] http://www.sas.com/en_us/software/business-intelligence/visual-analytics.html
(Link posećen: 01.11.2014.)
- [26] <http://www.knewton.com/> (Link posećen: 15.03.2015.)
- [27] <http://www.aleks.com/> (Link posećen: 15.03.2015.)
- [28] <http://uki.blackboard.com/sites/international/globalmaster/Platforms/>
(Link posećen: 15.03.2015.)
- [29] <http://www.ellucian.com/> (Link posećen: 15.03.2015.)
- [30] <http://dmlcentral.net/blog/lyndsay-grant/understanding-education-through-big-data>
(Link posećen: 15.03.2015.)
- [31] Weiyi Shang, Zhen Ming Jiang, Hadi Hemmati, Bram Adams, Ahmed E. Hassan, Patrick Martin, "Assisting Developers of Big Data Analytics Applications When Deploying on Hadoop Clouds" - 978-1-4673-3076-3/13 2013 IEEE, , San Francisco, CA, USA, ICSE 2013
- [32] „Worldwide File-Based Storage research“, IDC, 2013
- [33] <http://www.blackboard.com/Platforms/Analytics/Products/Blackboard-Analytics-for-Learn.aspx>
(Link posećen: 01.07.2014.)
- [34] http://www.ted.com/talks/daphne_koller_what_we_re_learning_from_online_education.html
(Link posećen: 15.06.2014.)
- [35] <http://www.dreambox.com/> (Link posećen: 22.03.2015.)
- [36] <http://hackeducation.com/2013/10/17/student-data-is-the-new-oil/> (Link posećen: 20.11.2015.)

- [37] <http://www.forbes.com/pictures/lmm45emkh/7-hod-lipson-and-michael-schmidt-computer-scientists-cornell-university/> (Link posećen: 22.06.2014.)
- [38] <http://www.skladistenje.com/big-data-u-sportu/> (Link posećen: 22.06.2014.)
- [39] <http://www.stats.com/sportvu/sportvu.asp> (Link posećen: 14.02.2015.)
- [40] <http://www.skladistenje.com/trendovi-u-zdravstvu/> (Link posećen: 15.06.2014.)
- [41] <https://www.touchsurgery.com/analysis/big-data-in-healthcare/> (Link posećen: 18.11.2016.)
- [42] <http://www.dummies.com/how-to/content/the-importance-of-virtualization-to-big-data.html>
(poslednja poseta 08.07.2014)
- [43] <http://www.forbes.com/pictures/lmm45emkh/3-sebastian-thrun-and-peter-norvig-data-scientists-google/> (Link posećen: 01.08.2014.)
- [44] <http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/> (Link posećen: 01.07.2014.)
- [45] Joel Hartman, Welcome to the Era of Big data and Predictive Analytics in Higher Education - University of Central Florida, 2013
- [46] <http://hackeducation.com/2012/12/09/top-ed-tech-trends-of-2012-education-data-and-learning-analytics/> (Link posećen: 18.10.2015.)
- [47] <http://augmentedtrader.wordpress.com/2013/05/25/mooc-student-demographics-fall-2012/>
(Link posećen: 22.06.2014.)
- [48] Dina Crnovršanin, MOOC (Otvoreni masovni onlajn kursevi), Fakultet tehničkih nauka, Čačak 2013
- [49] <http://www.edtechmagazine.com/higher/article/2014/02/comparison-five-free-mooc-platforms-educators> (Link posećen: 15.06.2014.)
- [50] Daphne Koller, What we're learning from online education, TedGlobal, 2012
- [51] <http://www.skolskidnevnik.net/2013/02/drustvene-mreze-u-nastavi/>
(Link posećen: 10.07.2014.)

- [52] Ana Miletić, Vesna Petković, Slobodanka Đenić, Verica Vasiljević, Mirko Stupar, Korišćenje alata Gephi na predmetu Društvene mreže, Visoka škola elektrotehnike i računarstva strukovnih studija, Beograd, Srbija, 2014
- [53] <http://www.adriatalk.com/drustvene-mreze-u-srbiji/> (Link posećen: 12.04.2015.)
- [54] <https://Twitter.com/edmodo/status/591766822245900289/> (Link posećen: 12.04.2015.)
- [55] Kelly Ryan, Twitter Study Reveals Interesting Results About Usage 40% is 'Pointless Babble', Pear Analytics blog post, 2014
- [56] Roosevelt C. Mosley Jr., Social Media Analytics: Data Mining Applied to Insurance Twitter Posts, FCAS, MAAA, 2014
- [57] David John Hughes, A tale of two sites: Twitter vs. Facebook and the personality predictors of social media usage, Manchester Business School East, The University of Manchester, 2015
- [58] Paul C. Zikopoulos, Dirk deRoos, Krishnan Parasuraman, Thomas Deutsch, David Corrigan, James Giles - „Harness the Power of Big data“, McGrawHill, 2012
- [59] "Data, data everywhere" - *The Economist*, Retrieved 9 December 2012.
- [60] Anirban Mukherjee, Joydip Datta, Raghavendra Jorapur, Ravi Singhvi, Saurav Haloi, Wasim Akram, Shared Disk Big data Analytics with Apache Hadoop, Symantec Corporation ICON, Baner Road, Pune – 411021, India. 978-1-4673-2371-0/12 - 2012 IEEE
- [61] <http://simplymeasured.com/> (Link posećen: 13.04.2015.)
- [62] Rastko Pavlović, Rastko Dejanović, Big Data i poslovna inteligencija, Infoteh-Jahorina Vol.13, Mart 2014
- [63] Dragoš Cvetković, Akademska misao, Spektralna teorija grafova i kombinatorna teorija matrica, Beograd, 2012
- [64] Popović M., Stanković R., Milovanović I., Milovanović E., Some Remarks on Normalized Laplacian Spreads of Graphs, Revista Tecnica De La Facultad De Ingenieria Universidad Del Zulia; ISSN: 0254-0770, Facultad de Ingeniería, Universidad del Zulia, Vol. 38, N° 2, pp. 13-17, 2015

[65] Milovanović E., Popović M., Stanković R., Milovanović I., Remark on ordinary and Randić energy of graph, Journal of Mathematical Inequalities, dx.doi.org/10.7153/jmi-10-55, pp. 687-692, 2015

[66] Milovanović I., Milovanović E., Popović M., Stanković R., Remark on the Laplacian-energy-like and Laplacian incidence energy invariants of graphs, Creative Mathematics and Informatics, Ms. Ref. No.: CMI-15-01303R1, ISSN 1843 - 441X, No. 2, pp. 183-187, 2015

[67] Seth A. Myers, Aneesh Sharma, Pankaj Gupta, Jimmy Lin, Information Network or Social Network? The Structure of the Twitter Follow Graph, WWW'14, 2014, Seoul, Korea, ACM 978-1-4503-2745-9/14/04

[68] <http://www-01.ibm.com/software/data/infosphere/biginsights/quick-start/downloads.html>

(Link posećen: 20.10.2015.)

[69] <https://dev.Twitter.com/start> (Link posećen: 22.10.2015.)

[70] <http://blog.cloudera.com/blog/2012/09/analyzing-Twitter-data-with-hadoop/>

(Link posećen: 02.11.2015.)

[71] <https://immersion.media.mit.edu/> (Link posećen: 03.11.2015.)

[72] <https://goo.gl/W57Qaj> (Link posećen: 03.11.2015.)

[73] <https://litmus.com/email-analytics>

Indeks slika

Slika 1. Veb konferencija	6
Slika 2. Primeri mobilnog obrazovanja	9
Slika 3. Popularan izgled Android operativnog sistema.....	11
Slika 4. Second Life u edukaciji.....	13
Slika 5. Prikaz Second Life učionica	14
Slika 6. Vežbe u virtualnoj ambulanti.....	15
Slika 7. Open Wonderland platforma [11]	16
Slika 8. Open Wonderland učionica	17
Slika 9. Jedna od definicija Big data.....	19
Slika 10. Svakog minuta se generišu ogromni podaci [15]	21
Slika 11. Big data ekosistem	23
Slika 12. 3D Vizualizacija - ITscape [22]	28
Slika 13. Polymaps vizualizacija [23].....	28
Slika 14. Q Research Software [24]	29
Slika 15. SAS Visual Analytics [25].....	30
Slika 16. Popularna ilustracija korišćenja Big data [30].....	31
Slika 17. Blackboard sveobuhvatni izveštaj [33]	33
Slika 18. Aplikacija za prikaz statistike NBA igrača.....	35
Slika 19. Big data u medicini [41].....	36
Slika 20. Zanimljiva ilustracija povodom ugrožavanja privatnosti studenata [45]	38
Slika 21. Ilustracija vrednosti informacija [46].....	39
Slika 22. Uopšteno i ukratko poređenje još nekoliko MOOC platformi.....	43
Slika 23. Društvene mreže.....	45
Slika 24. Youtube u edukaciji	53
Slika 25. Društvena mreža Edmodo – kreiranje profila.....	55
Slika 26. Wikipedija u edukaciji.....	56
Slika 27. Izgled ponude za analizu društvenih mreža kompanije „Simply Measured“ [61].....	58
Slika 28. Twitter graf.....	62
Slika 29. Primena NeoGraph simulatora na Twitter društvenu mrežu i određivanje Laplacian energije	64
Slika 30. Zanimljiv slajd sa IBM-ove prezentacije o Big data	66
Slika 31. Izgled CentOS nakon uspešne instalacije	67
Slika 32. Startovanje instalacije IBM BigInsights	68
Slika 33. Prikaz završene instalacije IBM BigInsights	69
Slika 34. Kreiranje zahteva za analizu u Twitter aplikaciji	70
Slika 35. cURL komanda potrebna za kreiranje „json“ fajla.....	71
Slika 36. Pokretanje cURL command-e	71
Slika 37. Izgled terminala nakon završetka.....	72
Slika 38. Izgled Eclipse prozora nakon uspešno izvršenih podešavanja	75
Slika 39. Početna stranica IBM BigInsights.....	76
Slika 40. Učitavanje „and.dell“ fajla u CSV formatu	77
Slika 41. Funkcija count	80
Slika 42. Izgled dobijene tabele.....	80

Slika 43. Vizuelni prikaz kojim jezikom se najviše tvitovalo za pojam "android"	81
Slika 44. Prikaz kako se neke komponente mogu složiti [70].....	83
Slika 45. Prikaz sentimentalne analize	84
Slika 46. Vizuelni prikaz analize email-a	86
Slika 47. Detaljan prikaz korisnika od koga je primljeno najviše mail-ova	86
Slika 48. Pokretanje Word Count aplikacije	88
Slika 49. Dodavanje funkcije "URL"	91
Slika 50. Sortirani linkovi iz tvitova koji su najpopularniji	92
Slika 51. Link koji se najviše svideo korisnicima u tom momentu	93
Slika 52. Registracija Twitter Service Provider funkcije.....	95
Slika 53. Upotreba tokena	95
Slika 54. Prosleđivanje parametara \$query	96
Slika 55. Grafički interfejs za prikupljanje podataka.....	96
Slika 56. Izgled aplikacije nakon popunjenih zahteva za analizu podataka.....	97
Slika 57. Rezultati nakon upita	98
Slika 58. Statistički prikaz rezultata prvog upitnika.....	106
Slika 59. Statistički prikaz rezultata drugog upitnika.....	110

Indeks tabela

<i>Tabela 1. Razlika između relacionih baza podataka i MapReduce</i>	<i>25</i>
<i>Tabela 2. Poređenje platformi Coursera, Udacity i Edx</i>	<i>42</i>
<i>Tabela 3. Odgovori anketiranih korisnika na pitanje br 1.....</i>	<i>47</i>
<i>Tabela 4. Odgovori anketiranih korisnika na pitanje br 2.....</i>	<i>48</i>
<i>Tabela 5. Značenje kolona u IBM BigInsights izveštaju</i>	<i>79</i>
<i>Tabela 6. Različiti upiti za Twitter aplikaciju.....</i>	<i>89</i>

DDR - ANALIZA TWITTER PODATAKA I MOGUĆNOSTI PRIMENE U EDUKACIJI

ORIGINALITY REPORT

4%

SIMILARITY INDEX

PRIMARY SOURCES

1	cris.uns.ac.rs Internet	213 words — 1%
2	aamirafridi.com Internet	144 words — 1%
3	www.havlena.net Internet	135 words — 1%
4	www.valentinkuleto.com Internet	91 words — < 1%
5	www.adriatalk.com Internet	63 words — < 1%
6	www.singipedia.com Internet	55 words — < 1%
7	infoteh.etf.unssa.rs.ba Internet	50 words — < 1%
8	npao.ni.ac.rs Internet	23 words — < 1%
9	bib.irb.hr Internet	19 words — < 1%
10	tjfeonline.com Internet	18 words — < 1%
11	www.itdesk.info Internet	18 words — < 1%
12	obuka.fon.rs	

	Internet	17 words — < 1%
13	www.singipedia.singidunum.ac.rs Internet	17 words — < 1%
14	www.pdf-flip.com Internet	16 words — < 1%
15	dev.twitter.com Internet	12 words — < 1%
16	uvidok.rcub.bg.ac.rs Internet	11 words — < 1%
17	"TWITTER DATA ANALYTICS IN EDUCATION USING IBM INFOSPHERE BIGINSIGHTS", Proceedings of the International Scientific Conference - Sinteza 2016, 2016. Crossref	9 words — < 1%
18	www.savezenergeticara.org.rs Internet	9 words — < 1%
19	b92.net Internet	9 words — < 1%
20	www.techpg.net Internet	8 words — < 1%
21	www.dms.org.rs Internet	8 words — < 1%
22	www.radista.info Internet	8 words — < 1%
23	issuu.com Internet	8 words — < 1%
24	itnetwork.rs Internet	8 words — < 1%
25	www.jrga.org Internet	

8 words — < 1%

EXCLUDE QUOTES OFF
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES < 5 WORDS