

UDK 657.6:004.6

PREGLEDNI RAD

DOI: 10.7251/FIN1803022M

Ljubiša Mičić*

Big Data i upotreba u savremenom poslovanju

Big Data and its usage in modern business

Rezime

Evidentan je globalni rast veličine podataka. Obim i složenost prikupljenih podataka eksponencijalno se povećavaju.

Povećanje obima podataka dovelo je do situacije u kojoj klasični sistemi za sakupljanje podataka, skladištenje, obradu i vizuelizaciju podataka gube bitku sa velikom količinom podataka, brzinom i raznovrsnošću podataka koji se kontinuirano generišu. Ovo generisanje podataka dovelo je do novih metodologija i tehnika skladištenja, pretraživanja, obrade i menadžmenta podacima, koji su objedinjeni kao nova naučna disciplina pod nazivom Big Data. Primjena Big Data je višestruka, od ekonomskih djelatnosti kao što su marketing i prodaja, do specifičnih djelatnosti kao što su saobraćaj i medicina. Tako je ovaj pristup pronašao svoju primjenu i u različitim segmentima poslovanja, od prodaje i nabavke do finansijske revizije.

Rad kroz pregled stručne i naučne literature objavljene na ovu temu analizira ključna pitanja kao što su da li i u kom obimu pristup Big data može pomoći u identifikaciji uzročno-posljedičnih veza između poslovnim podacima, zakonitosti kao i šema ponašanja; koje oblike koristi imamo na osnovu velikog seta podataka; kako ti iste koristi mogu biti integrisane u tradicionalne procese poslovanja, kao i kakve to efekte ima na sam kvalitet poslovanja. Rezultat istraživanja je pregled odgovora iz vodeće literature na gore postavljena pitanja.

Ključne riječi: Big Data, podaci, poslovanje.

Abstract

There is a global increase of data volume. Volume and complicity of data is exponentially increasing which has led to situation in which traditional database management systems, as well as those for data visualization are losing game with big data volume, speed and diversity which are continuously being generated. This process has let to new methodologies and techniques of data warehouse, data search as well as data management integrated to new scientific discipline Big Data, Big data application is diverse, from business activities such are marketing and sales to specific ones such are traffic and medicine. Therefore, this concept has been used in different segments from sales to financial auditing. Paper presents review of researches published on this subject and provides answers to following questions: can Big data helps in identification of causes and outputs of different business tasks, legality and schemes of behavior; which usage can be obtained from large sets of data; how these things can be integrated in traditional business processes and which effects does it have on quality of doing business. Result of research is review of answers from leading sources connected to above mentioned questions.

Keywords: Big Data, data, business.

* Univerzitet u Banjoj Luci, Ekonomski fakultet, e-mail: ljubisa.micic@ef.unibl.org, Univerzitet u Novom Sadu, Ekonomski fakultet Subotica, doktorand u oblasti poslovne informatike.

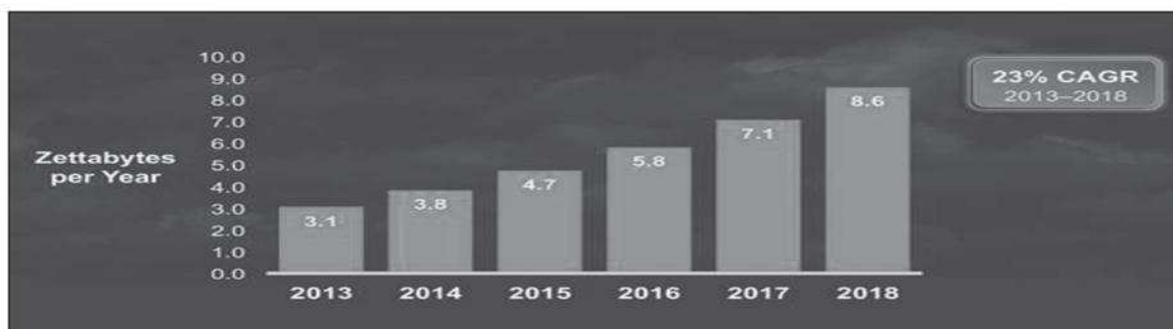
UVOD

Količina podataka na svjetskom nivou eksponencijalno raste. Uporedo s ovom pojavom pojavila se potreba za novim jedinicama mjere, pa su pored gibabajta, terabajta, petabajta uvedeni eksabajt, zetabajt i jotabajt kao posljednja jedinica mjere količine podataka. Bitno je napomenuti da je svaka naredna jedinica mjere jedna hiljada prethodne tako da je na kraju jotabajt 10^{24} bajtova. Rast količine podataka doveo je do situacije da klasični sistemi za prikupljanje, skladištenje, obradu i vizuelizaciju podataka gube bitku sa velikom količinom, brzinom i raznovrsnošću podataka koji se neprestano generišu, a koji su nazvani Big Data. Sve je više podataka koji su kreirani od strane Internet of Things, IoT (kamere, sateliti, automobili, GPS navigacije i dr.). Na nama je izazov da osmislimo nove tehnologije i alate za upravljanje i eksploataciju tih velikih količina podataka. Big Data je veoma aktualna tema posljednjih godina u IT krugovima. Međutim, Big Data je doživio veliku afirmaciju i u poslovnom svijetu, a sve više i u javnoj upravi. Potencijal eksploatacije Big Data u svim aspektima ljudskog djelovanja i rada je sve više počeo da zanima i IT stručnjake, ali i osobe iz poslovnog svijeta u zemljama u okruženju (Hrvatska, Srbija, Slovenija). Ovaj rad se odnosi na problematiku koja obrađuje upotrebu Big Data u privredi. Predmet rada predstavlja analiza potencijala koje pruža Big Data za različite segmente poslovanja i načina njihovog iskorištavanja. Cilj rada je da se ukaže na dalji razvoj Big Data, odnosno usmjeri na buduće pravce razvoja i da doprinos sa teoretskog aspekta Big Data, tehnologijama Big Data, potencijalima i daljim pravcima razvoja, kao i njihov uticaj na savremeno poslovanje.

1. DEFINISANJE BIG DATA

Da bi se jasno definisao pojam Big Data (BD), prvo je neophodno definisati sam podatak i njegovo razlikovanje od informacije. Postoje različiti pogledi u definisanju pojmova „podaci“ i „informacije“. Jedan pogled zastupa tezu da su podaci strukturirane činjenice na način koji je pogodan za obradu ljudima i računarima, a koje

Slika 1. Porast količine podataka na svjetskom nivou u periodu 2013–2018. u ZB



Izvor: Cisco Global Cloud Index, 2013–2018. (Cisco, 2013)

Danas, kao glavni izvori rasta podataka mogu se identifikovati: a) kartice lojalnosti klijenata za trgovačke centre, koje predstavljaju izvor hronoloških podataka o mjestima kupovine, kupovnim navikama i kupovnoj moći potrošača – utiču na stalni rast baza podataka o klijentima (CRM – Customer Relationship Management); b) podaci iz oblasti medicine i zdravstva kao što su e-kartoni pacijenata, recepti, snimci operacija, zapisi raznih vrsta pregleda, analiza i pretraga; c) finansijski podaci sa svjetskih, regionalnih i lokalnih berzi, podaci firmi i državnih regulatornih tijela; d) društvene mreže doprinose „bumu“ podataka, jer korisnici sami postavljaju ne jednu, već veći broj fotografija i drugih sadržaja dnevno; e) Internet of Things (IoT), kada uređaji (mašine) komuniciraju, generišu podatke, npr. u

nemaju značenje. Podatak postaje informacija tek kad poprimi značenje (Aleksić-Marić, 2008). Npr., podatak „38“ sam za sebe nema određeno značenje, ali ako je to ukupan broj studenata na master studiju određenog fakulteta – onda navedeni podatak postaje informacija. Drugi rasprostranjeniji pogled zastupa tezu da podaci postaju činjenice tek onog momenta kada su one od važnosti za nas. Tako, npr., činjenica „Očekivana temperatura u Tel Avivu za 8.8.2018. je 39° C“ za nas je informacija tek onda kada nam je od značaja, npr. ukoliko planiramo put u Tel Aviv tog datuma ili npr. ukoliko radimo istraživanje poređenja temperatura u tom gradu u određenom periodu.

1.1. Nastanak Big Data

Količina generisanih podataka u svijetu kontinuirano raste. Prema predviđanjima Ciscoa (Cisco, 2013), ukupan broj generisanih podataka u svijetu krajem 2015. iznosiće oko 4,7 zetabajta (ZB). Od početka vremena, do kraja 2003. godine, ukupno je generisano 5 EB podataka za cijelo čovječanstvo, a od 2004. do kraja 2012. čak 2,7 ZB elektronskih podataka (slika 1), dakle, u periodu od devet godina generisano je 500 puta više podataka nego za sav protekli istorijski period (Intel, 2013). Progresija podataka se može posmatrati kroz prizmu kreatora podataka. Prve podatke su unosili zaposleni u kompanijama, pojavom Weba 2.0 (prije njega korisnici nisu unosili sadržaje jer je sadržaj bio preglednog karaktera a ne interaktivnog tipa). Korisnici su ti koji unose dominantnu količinu podataka u odnosu na tradicionalne kreatore podataka. Posljednja grupa kreatora podataka, koja danas dominira u odnosu na preostale kreatore, jesu mašine, odnosno uređaji – IoT (Internet of Things). Podaci rastu eksponencijalnom brzinom te se ne stignu sačuvati prije analize i sređivanja. Nagli rast podataka je izazvan velikim dijelom od strane uređaja, jer oni sve više generišu podatke, komuniciraju međusobno, ali i sa ljudima. Tradicionalna tehnologija koja je obrađivala malu količinu podataka zasnivala se na sljedećem procesu: podaci bi bili sakupljeni i dostavljani na obradu CPU (Central Processing Unit), međutim, vremenom je količina podataka dostigla kolosalni nivo i CPU više nisu mogle obrađivati toliku količinu podataka. Nova tehnologija temelji se na obrnutom procesu, više CPU se „donosi“ na izvor podataka i dolazi do tzv. paralelne obrade (više podataka se obrađuje istovremeno).

laboratorijama, proizvodnji, domaćinstvima, saobraćaju, bezbjednosti itd. Svi ovi podaci se nalaze u bazama podataka koje konstantno rastu. Takođe, povećan broj tableta, pametnih telefona, pametnih satova i drugih uređaja omogućava pristup internetu sa svakog mjesta i u svako vrijeme, čime rastu cloud baze podataka; f) novi oblici podataka koje generišu naučnici – npr. podaci o strukturi i sastavu DNK koji su veoma veliki i kompleksni, predstavljaju novo polje na kome naučnici sprovode ispitivanja.

1.2. Definicije osnovnih domenskih konceptata

Big Data predstavlja poseban koncept digitalnih informacija koje su nepogodne za skladištenje, transport i/ili analizu tradicionalnim

pristupima. Ovi podaci su toliko obimni da prevazilaze tehnološke mogućnosti za skladištenje, analizu i vizuelizaciju današnjice (Mir, 2013). Izazov je kreiranje naredne generacije posebnih tehnika za skladištenje, transport i analiziranje ovih podataka.

Problem velikih podataka nije naročito novi koncept. Sa ovim problemom su se prvi put susreli fizičari u CERN-u, Švajcarska, prije pedesetak godina. Očigledan primjer poteškoća u upravljanju i analiziranju velikih podataka na koji su naišli bio je na projektu otkrivanja Higgsovog bozona, 2012, jer detektori, kojih ima 1.050 miliona, hvataju tragove čestice zumiranjem protona koji se kreću brzinama svjetlosti u akceleratoru, 40 miliona puta svake sekunde, što predstavlja ozbilju količinu podataka koju treba analizirati (Smith, 2013).

1.3. Kad Data postaje Big Data?

Rastom količine, brzine pristizanja i nestrukturisanosti podataka tradicionalni sistemi za obradu podataka gube na svom značaju. Npr. ukoliko se posmatra dokument veličine 100 MB, ne može se poslati mailom, slika veličine 100 GB ne može se vidjeti, a video-zapis od 100 TB ne može se uređivati tradicionalnim alatima. Alati nove generacije, Big Data Tools, morali bi riješiti ova, ali i mnoga druga pitanja. Područja od posebnog značaja za alate nove generacije su (Mir, 2013):

Tabela 1. Primjeri nastanka Big Data

Događaj	Opis
Let aviona Air Bus	10 TB se generiše na svakih 30 minuta leta, tokom jednog leta prosječno nastane 640 TB podataka.
Digitalna brojila potrošnje struke (smart meters)	U 2009. je bilo instalirano oko 76 miliona digitalnih brojila električne energije, što je rezultovalo proizvodnjom 350 biliona transakcija godišnje. Procjena je da će postojećim tempom 2020. godine broj ovih uređaja iznositi 800 miliona.
Kamere na mobitelima	Tokom 2013. u svijetu je bilo oko 5 milijardi kamera na mobitelima. Većina njih ima GPS. Taj broj se iz dana u dana povećava, kao i broj fotografija i snimaka koje korisnici kreiraju.
Internet korisnici	Krajem 2014. broj internet korisnika je iznosio gotovo 3,08 milijardi, što je doprinijelo da se količina internet saobraćaja mjeri u ZB.
Blogovi	Podaci iz 2013. govore o broju od 200 miliona registrovanih blogova.
Elektronska pošta	Oko 300 milijardi mejlova se pošalje svaki dan.
RFID	Primjena RFID tagova (čipova) u porastu je na svjetskom nivou, 2005. je bilo 1,5 miliona tagova, a već sedam godina poslije – 30 milijardi tagova.
Twitter	Ova društvena mreža generiše oko 12 TB podataka dnevno, sa 200 miliona korisnika i oko 230 miliona twittova dnevno (97.000 twittova u sekundi).
NYSE	Njujorška berza dnevno generiše oko 1 TB podataka.

Izvor: Mir, 2013; Stats, 2015; Telefonica, 2014.

Krajem 2014. godine, ukupna količina generisanih podataka, prema procjenama sa početka rada, iznosila je oko 4,7 ZB. Nakon prikazanih izvora podataka, načina, rezultata i količine generisanja istih, slika o Big Data se usložnjava i dobija na sve većem značaju.

1.4. Karakteristike Big Data

Osnovne osobine BD mogu se predstaviti modelom 3V: volume (obim, odnosno količina), velocity (brzina), variety (raznovrsnost). Veliki obim, odnosno velika količina podataka je prva osobina BD i predstavlja izazov, ali i priliku za velike organizacije da bolje shvate ljudsko ponašanje i shodno tome reaguju, odnosno prilagode svoje resurse. Primjer Facebooka bi na ilustrativan način mogao doprinijeti boljem razumijevanju velikog obima podataka koji svake minute pristižu u bazu podataka. Korisnici Facebooka svojim aktivnostima i korištenjem usluga koje nudi ova društvena mreža, stvaraju 25 TB podataka dnevno (1,04 TB svaki sat). Brzina kojom pristižu velike

- 1) prikupljanje,
- 2) sadržaj,
- 3) skladištenje,
- 4) analiziranje,
- 5) pretraga,
- 6) dijeljenje,
- 7) transfer i
- 8) vizuelizacija podataka.

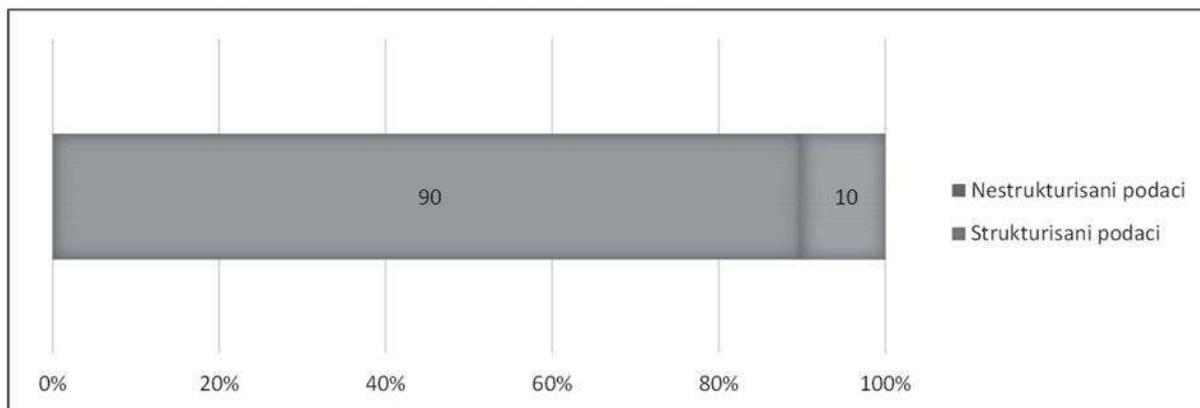
Kad neki podatak postaje BD zavisi od kapaciteta sistema organizacije. Što za jednu organizaciju može biti BD, za drugu ne mora. Mir smatra da podaci postaju BD kad dostignu veličinu od 1 TB (Mir, 2013). Primjeri tačaka generisanja podataka i tipova podataka prikazani su slikom 2. Veliki broj različitih vrsta uređaja postaje smart i odašilje podatke. Podacima koji nastaju su immanentni brzina, velika količina i raznovrsnost (nestrukturisanost).

Kako bi se bolje prikazale razmjere nastanka BD, u tabeli 1. je navedeno nekoliko primjera događaja i rezultat istih u pogledu nastale količine podataka.

količine podataka je usko povezana sa kapacitetima IT sistema i načinima eksploatacije BD u realnom vremenu (Fast Data). Ova osobina BD otvara sasvim novo polje za istraživanje. Primjer pristizanja velike količine podataka je švajcarski CERN, u kome, prilikom obavljanja ispitivanja, svake sekunde pristiže 40 TB podataka.

1.5. Raznovrsnost podataka

Podaci dolaze velikom brzinom, u velikom obimu, ali je, pored toga, prisutna i velika raznovrsnost njihovih oblika. Podatke možemo podijeliti na strukturisane (tekstualni podaci – stringovi, numerički, dokumenti, finansijski podaci, zapisi, lični podaci i dr.) i nestrukturisane (fotografije, tekstualne poruke, audio i video zapisi, 3D modeli, simulacije, podaci o geografskoj lokaciji i sl.). Danas imamo mnogo nestrukturisanih podataka, što ih čini nepodobnim za tradicionalni način obrade podataka (slika 3).

Slika 2. Struktura informacija pri donošenju odluka na primjeru poslovnog preduzeća

Izvor: Adaptirano prema: Mir, 2013.

Ovo je veoma bitno područje prilikom analiziranja procesa donošenja odluka. Npr. ukoliko se posmatra jedan privredni subjekat koji posluje na tržištu Evropske unije, struktura podataka koje ova firma koristi za odlučivanje (predstavljena slikom 3) sastoji se od strukturisanih (podaci iz finansijskih i nefinansijskih izvještaja preduzeća) i nestrukturisanih podataka (koji takođe nastaju u preduzeću, ali dolaze i izvan njega). Od ukupno raspoloživih podataka za donošenje odluka, firma koristi svega 10%. Preostalih 90% podataka firma ne koristi pri donošenju odluka i upravo u ovom području treba tražiti priliku za rast i razvoj. U slučaju da preduzeće može prikupiti sve nestrukturisane podatke, oni obično ne bivaju analizirani, a u slučajevima u kojima se radi na njihovoj analizi – ne dobijaju se vrijednosti, odnosno znanje ili informacije od vrijednosti za poslovanje, jer tradicionalni sistemi nisu u stanju analizirati ove podatke i ispostaviti informacije od vrijednosti za firmu (Mir, 2013). IBM-ova definicija BD obuhvata 4V (volume, velocity, variety, veracity). Posljednji termin se odnosi na (ne)pouzdanost podataka, odnosno njihovu istinitost. U IBM-u se ovo ističe kao veliki problem, jer 3,1 milijardi dolara čine troškovi koji nastanu u USA na godišnjem nivou zbog netačnih podataka (IBM Big Data Hub, 2015).

2. TEHNOLOGIJA BIG DATA

Zbog ograničenja u tehnologiji hardvera, podaci danas više ne mogu fizički da stanu na jedan disk ili USB, zbog čega se moraju distribuisati (rasporediti, raspodijeliti) na više mjesta. Kada su podaci distribuisani, sljedeći korak je paralelna obrada podataka. Paralelna obrada podataka je situacija u kojoj više procesora paralelno radi na izvršenju određenog zadatka koji je postavio klijent. Ovakav način obrade podataka dovodi do ušteda u troškovima i vremenu, dakle, povećava efikasnost obrade.

Ukoliko pokušamo da sačuvamo dokument od 1 TB, ili da ga otvorimo, analiziramo, uređujemo – za to će nam trebati mnogo vremena ukoliko koristimo standardne alate. Big Data tehnologija, stoga, treba da omogući brzo skladištenje, pristup i analizu velikih fajlova. Pored toga, BD tehnologije moraju da ispoštuju određena načela i principe poslovanja koji su prikazani u tabeli 2.

Tabela 2. Faktori od uticaja na BD tehnologiju i komparativnu prednost organizacije

Faktor	Uloga
Sistemi	Pred infrastrukturu se postavlja zahtjev eksploatacije podataka u realnom vremenu (instantno). Sistem bi trebalo da bude optimizovan na način da u realnom vremenu (ili blizu toga) automatski odgovara na zahtjeve poslovnih procesa.
Pohranjivanje/čuvanje podataka	Za postizanje efektivnosti i efikasnosti potrebno je imati aktivne analitičke procese dok podaci „teku“ kroz organizaciju. Organizacija bi trebalo da definiše koji podaci se čuvaju i u skladu s tom politikom kreira infrastrukturu koja će zadovoljiti potrebe za čuvanjem, smanjenjem troškova i potrebe u pogledu redukcije rizika gubitka podataka.
Načelo povjerljivosti	Da bi organizacija sebi osigurala dobar poslovni rejting, neophodno je da u svojim poslovnim procesima ima inkorporirane rigidne politike i prakse u vezi sa povjerljivosti i zaštitom podataka na kojima bazira svoje poslovanje.
Načelo sigurnosti	Načelo sigurnosti podataka treba da bude u središtu pažnje organizacije, na čemu ona može da gradi svoju konkurentnu prednost. Infrastruktura organizacije bi trebalo da ima značajne sigurnosne sisteme kako bi podaci i organizacija bili zaštićeni od internih i eksternih prijetnji.
Upravljanje	Odgovarajuća platforma uliva povjerenje, pa se kompanija može osloniti na istu. Ona kontroliše kreiranje podataka, dijeljenje, pročišćavanje, konsolidovanje, zaštitu, upravljanje, povlačenje i integraciju podataka.
Cloud	Kako bi organizacija smanjila pritisak koji BD pravi na njenu IT infrastrukturu, može donijeti odluku da pređe na cloud, odnosno da izmjesti BD i analitička rješenja. Time se postiže fleksibilnost, mogućnost proširenja i ekonomično angažovanje resursa koji će, svi zajedno, u budućnosti dovesti do povećanja konkurentne prednosti organizacije.

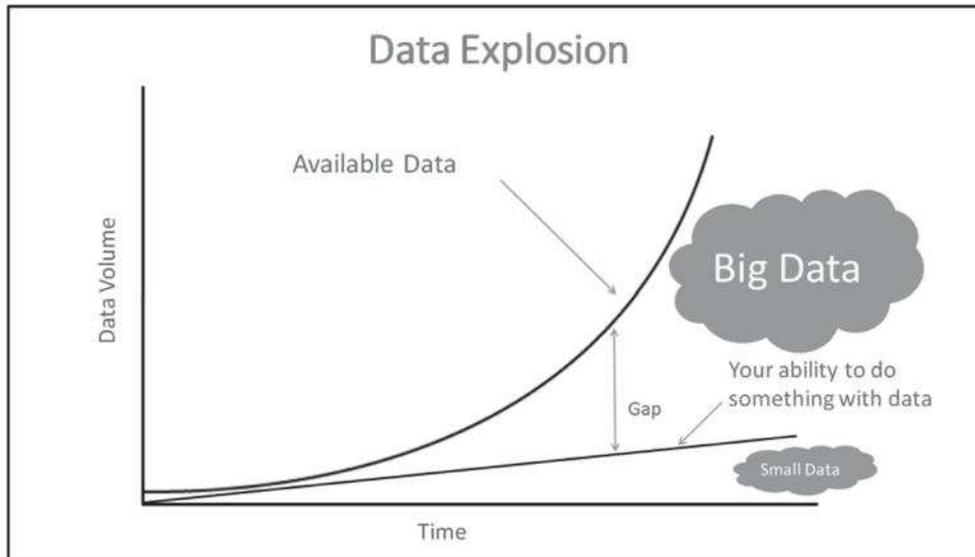
Izvor: IBM, 2014.

3. BIG DATA POTENCIJALI

Progresijom podataka tokom vremena, organizacije upadaju u svejevrni informacijski jaz, jer resursi kojima raspolažu mogu obraditi strukturisane (donekle i polustrukturisane) podatke manjeg obima (Small Data). Obim nestrukturisanih dostupnih podataka daleko je

veći i predstavljen je gornjom linijom. Prostor između podataka koje organizacije mogu da iskoriste (Small Data) i gornje linije na slici 3, koja predstavlja ukupnu raspoloživu količinu podataka koje organizacije ne iskorištavaju, predstavlja Big Data Gap, odnosno obim podataka koje organizacije ne uzimaju u obzir prilikom svog djelovanja.

Slika 3. Jaz u iskorištavanju podataka



Izvor: Mariner, 2014.

Prikazani jaz podataka predstavlja potencijal za organizacije kako bi ostvarile svoju konkurentsku prednost u odnosu na druge. Pema istraživanjima McKinsey Global Institute (McKinsey Global Institute, 2011), kompanije koje prepoznaju potencijal BD prilikom donošenja odluka i putem alata, kao što su Hadoop, nauče da dobijaju informacije i znanje – ostvariće konkurentsku prednost na tržištu za oko 20% u odnosu na kompanije koje ne koriste BD.

3.1. Primjeri iskorištavanja Big Data potencijala

Amazon

Ova kompanija našla je način kako da izvuče „znanje“ iz velikih količina različitih podataka kojima raspolaže. Amazon želi da sazna šta bi kupac (još) trebalo da kupi. Odgovor na ovo pitanje krije se u podacima o kupcu koje ima na osnovu: pretrage klijenta na sajtu, kretanja i zadržavanja po proizvodima, prethodne kupovine, a tu su i ostale informacije, kao što su geografska lokacija, podaci i aktivnosti na društvenim mrežama (preferencije kupaca od muzike, softvera, preko garderobe, putovanja i dr., iz koje je države, kojoj etničkoj zajednici pripada i sl.) Iz svih ovih podataka putem složenih algoritama koji se pokreću svaki put kad se kupac loguje na sajt – porede se aktivnosti svakog kupca u odnosu na drugog, čime se dobija odgovor na pitanje šta bi kupac trebalo (još) da kupi. Tako Amazon ima mogućnost, u realnom vremenu, dakle, dok kupac „razgleda“ po sajtu, da mu ponudi proizvode/usluge za koje je dobio rezultat da su u domenu interesovanja tog određenog kupca. Konkretni primjer bi izgledao ovako: kupac se loguje na sajt Amazona, čime pokrene algoritme za određivanje šta bi trebalo da kupi. Kupcu se na osnovu preferencija nudi knjiga X. Ili, ako je kupac kupio kupaći kostim, putem algoritma, ponudiće mu se obuća za plažu, ili neki drugi dodatak. Takođe, kupcu se može ponuditi garderoba na osnovu geografske lokacije, godišnjeg doba, pola, uzrasta i sl. (Mir, 2013).

Facebook

I ova kompanija je naučila da izvuče vrijedne informacije iz Big Data kojima raspolaže. Ciljevi su: 1) predložiti prijatelje, 2) targe-

tirati precizno publiku za oglas, 3) odrediti grupe za usmjeravanje određenih ponuda (npr. preporuka za knjige, muziku i sl.). Facebook ima pristup sljedećim podacima: a) prethodnim aktivnostima člana, b) trenutnim aktivnostima člana, c) aktivnostima drugih članova (na osnovu čega pravi komparacionu analizu) i d) drugim informacijama (starost, pol, geografska lokacija, preferencije u pogledu svega što je korisnik istakao i dr.). Facebook primjenjuje algoritme koji za cilj imaju da prepoznaju uzorke/šablone i svrstaju korisnike u različite segmente. Društvena mreža je postigla sljedeće: naučila je da iz velike količine podataka izvuče informacije i nova znanja kako bi ispunila svoja tri cilja i postigli su da sa velikom preciznošću za najveći broj korisnika mogu ispuniti ciljeve.

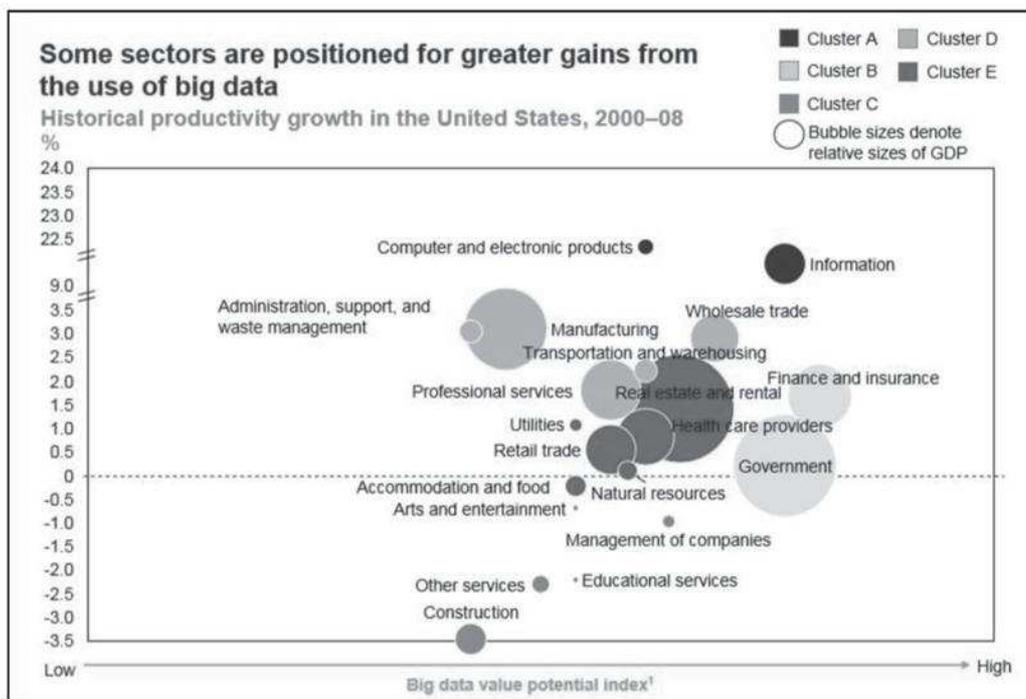
Google

Google je pokretač softverskih elemenata koji čine mogućim BD analizu. Tehnologiju je inicirao Google, a ona je prešla na Yahoo, a zatim na Apache. Google ima za cilj da iskoristi BD potencijale u punom kapacitetu. Cilj je da se 1) napravi procjena rezultata pretrage, 2) predvidi obim internet saobraćaja, 3) prikažu relevantni oglasi i 4) servisiranje klijenata putem Google aplikacija. Pošto je Google najpopularniji pretraživač (Vaughan-Nichols, 2015), već dugo vremena ima pristup informacijama iz web-pretrage, preferencijama korisnika, cookiesima, istoriji aktivnosti i dr. Algoritme koriste da na osnovu informacija koje posjeduju ostvare svoje ciljeve.

3.2. Oblasti sa najvećim mogućnostima za eksploataciju Big Data

U svakoj oblasti ljudskog djelovanja i rada danas su prisutni digitalni podaci. Osim za osobe iz oblasti IT-a, Big Data je postala veoma važna oblast i za top menadžment preduzeća i organizacija u mnogim sektorima. Ako se izvrši komparacija privrednih i neprivrednih oblasti, u nekima se može više, a u nekima manje iskoristiti potencijal BD.

Slika 4. Potencijali eksploatacije BD prema sektorima u SAD



Izvor: McKinsey Global Institute, 2011.

Prema istraživanju McKinsey Global Institute (McKinsey Global Institute, 2011), klaster A (koji čine kompjuterski i elektronski proizvodi, te informacioni sektor) imao je najveći procenat rasta produktivnosti (Y-osa) u posmatranom periodu, ali i najveće koristi od eksploatacije BD (X-osa) (slika 7, veličina krugova se odnosi na relativno učešće u GDP-u). Sektor koji je imao mali rast produktivnosti, a posjeduje značajne potencijale u oblasti BD je klaster B (finansije, osiguranje, upravljanje). Klaster C (konstrukcije, drugi servisi, obrazovanje, umjetnost, upravljanje kompanijama) doživio je pad produktivnosti koji se objašnjava jakim sistemskim barijerama koje sprečavaju rast produktivnosti. Klaster D (proizvodnja, administracija i upravljanje otpadom, transport i skladištenje i profesionalne usluge), koji je zastupljen na globalnom nivou, imao je značajan rast produktivnosti, a klaster E (nekretnine, zdravstvo, maloprodaja, prirodni resursi, komunalije), koji je više zastupljen na lokalnom nivou, nešto manji rast produktivnosti. Oba klastera imaju širok obim BD potencijala. Iako će svi sektori morati prevladati barijere koje im onemogućavaju eksploataciju BD, za neke će ovo biti više, a za druge manje zahtjevno. Npr. javni sektor i obrazovanje imaće više poteškoća

da stvore komparativnu prednost na osnovu BD, jer im nedostaje podataka i djelovanja kao organizacija koje su vođene podacima. Ograničene investicije i rast u zdravstvu su do sada doveli samo do relativno niskog nivoa performansi IT ulganja u zdravstvu. Sektori kao što su proizvodnja, prodaja, uslužni sektor mogu imati relativno male barijere upravo iz suprotnih razloga.

Ovaj izvještaj takođe ukazuje na mogućnosti i potencijale primjene tehnologija za određene sektore (tabela 3). Benefiti koji bi se ostvarili ogledali bi se u sljedećem: povećanje efikasnosti i efektivnosti, proizvodnja visokokvalitetnih proizvoda/usluga, odnosno stvaranje velike dodatne vrijednosti na proizvodima i uslugama. Npr. preduzetničke firme mogu uz upotrebu i analizu podataka kreirati dizajn proizvoda/usluga koji će ispuniti potrebe klijenata na većem nivou i/ili na bolji način. Podaci se mogu iskoristiti npr. da kreiraju dodatnu vrijednost proizvoda tokom same njegove upotrebe. Primjer za to je smart phone, koji „nauči“ potrebe i preferencije svog vlasnika, te uvijek drži spremne aplikacije i podatke koje vlasnik najčešće koristi, što za korisnika ima veću vrijednost nego potpuno novi uređaj koji nije kastomizovan.

Tabela 3. Procijenjeni benefiti u pet sektora eksploatacijom Big Data (2011)

Sektor	Procijenjena godišnja dodata vrijednost	% rasta produktivnosti na godišnjem nivou
Zdravstvo (SAD)	300 milijardi dolara	0,7%
Administracija, javni sektor (EU)	250 milijardi dolara	0,5%
Lični podaci o kretanju (globalni nivo)	100 milijardi dolara	700 milijardi krajnjim korisnicima direktno ili indirektno
Prodaja (SAD)	60% rast	0,5%–1,0%
Proizvodnja	50% rast	7% redukovani troškovi

Izvor: McKinsey Global Institute, 2011.

Nadalje, benefiti od ličnih podataka o lokaciji doprinose smanjenju troškova prevoza, goriva i uštedama u vremenu. To su sve načini generisanja dodatne vrijednosti na osnovu upotrebe velikih podataka. Prema procjenama, došlo je do značajnog rasta dodatne vrijednosti

u proizvodnoj, odnosno uslužnoj djelatnosti u pet gorenavedenih sektora kada su se više koristili podaci pri donošenju odluka. Najveći efekti su se osjetili u zdravstvu (u SAD), koje je eksploatacijom koristi od BD dobilo 300 milijardi dolara dodatne vrijednosti u 2011. godini

za svoje usluge, čime se podigao kvalitet za korisnike usluga. U Evropi je najviše benefita za korisnike usluga administracije javnog sektora, čak 250 milijardi dolara dodatne vrijednosti u toj godini, sa godišnjim povećanjem produktivnosti od 0,7% (tabela 3).

4. BIG DATA U REVIZIJI

Povezivanje poslovnih procesa i eksternih setova podataka dramatično mijenja poslovno okruženje. Pod jedan, organizacije koriste različite vrste cloud servisa (Weinman, 2013; Wei, 2014; Du, Cong, 2015). Ti servisi omogućavaju pristup različitim setovima Big Data podataka te samim tim generisanje različitih revizorskih dokaza. Dai (Dai, 2014) predlaže upotrebu Audit Data standarda (Zhang et al., 2012) kao sastavni dio integrisanog sistema revizije budućnosti. Standard predviđa da u svim koracima procesa softverski agenti generišu određeni oblik dokaza (Papazoglou, 2001). Bitno je napomenuti kako ova tehnologija može biti iskorištena da osigura da su atributi jasno definisani i standardi ispunjeni. Nekoliko bitnih stvari treba uzeti u razmatranje:

- Obim podataka nije primarna upotrebna funkcija. Savremena revizija će, osim obima, uzimati u obzir vremensku komponentu generisanih podataka kroz Big Data tehnologiju, tj. vremensku nerijetko trenutnu dostupnost velikog seta podataka za izvođenje revizorskih dokaza kroz različite analitičke alate.
- Sa aspekta kvaliteta generisanih podataka, bitnost i pouzdanost su ključne karakteristike upotrebe podataka u reviziji. Big Data značajno mijenja način posmatranja ove dvije karakteristike jer se ocjena bitnosti i pouzdanosti ne prepušta ljudskom faktoru nego složenim algoritmima zasnovanim na Big Data, nerijetko kroz automatizovane procese koju imaju veći stepen pouzdanosti od ljudskog faktora.
- Izvori i vrste revizorskih dokaza su novi i mogu zamijeniti do sada uobičajene tradicionalne dokaze, ali je potrebna dodatna analiza i istraživanja profesionalne struke u ovom domenu kako bi se izbjeglo ignorisanje grešaka iz automatizovanih dokaza.

ZAKLJUČAK

Big Data posljednjih godina doživljava punu ekspanziju. Veliki broj naučnih radova, izvještaja i projekata potvrđuje neophodnost uvođenja BD alata u svakodnevno poslovanje. Dok je ova oblast veoma popularna u svijetu, u zemljama regiona ona je tek u povoju. U Bosni i Hercegovini još uvijek ne postoje projekti uvođenja Big Data, niti forumi ili blogovi sa ovom tematikom. Takođe, Big Data minimalno se spominje u nastavnim planovima i programima u visokoškolskom obrazovanju na Univerzitetu u Banjoj Luci.

Informacioni jaz koji se javlja zbog nemogućnosti obrade nestrukturisanih podataka prepoznat je kao potencijal za organizacije koje premoštavanjem ovog jaza mogu da ostvare oko 20% konkurentne prednosti u odnosu na organizacije koje ne koriste BD alate. U radu su navedeni primjeri slučajeva iz prakse, gdje je na prvom mjestu najočigledniji primjer – Amazon, kompanija koja je naučila da iz mnoštva podataka (pretraživanje, klikovi po stranici, statusi na društvenim mrežama, istorija kupovine, kretanje po sajtu, mjesto stanovanja, klima, temperatura i dr.), uz pomoć algoritama, generiše znanje na osnovu kojeg predlaže kupcu artikle koje bi trebalo da kupi, ili koji bi mu bili interesantni. Slično je uradio i Facebook, samo što ova kompanija svoj BD koristi za targetiranje

klijenata za oglašavanje i marketing kampanje, putem veoma dobro definisanih filtera.

Privredne oblasti koje imaju najveći potencijal iskorištavanja BD u svrhu podizanja profitabilnosti, povećanja ekonomičnosti, odnosno smanjenja troškova su, na prvom mjestu, iz oblasti informacionog sektora, zatim slijede finansijski sektor, upravljanje, trgovina i javni sektor. Naravno, i revizija je jedna od djelatnosti gdje bi ova tehnologija mogla biti korištena.

Generalno, svi autori ističu informacioni jaz između iskorištenih podataka za donošenje odluka i ukupno dostupnih podataka, od kojih je ogromna većina nestrukturisana. Organizacije ignorišu kolosalne količine podataka prilikom donošenja odluka. BD daje preduzećima i organizacijama duboki uvid u podatke o poslovnim partnerima, potrošačima, vremenskim prognozama, pandemijama, poslovnim predviđanjima, saobraćaju i transportu i drugim oblastima.

Pravci daljih istraživanja bi mogli da se odnose na tehničke uslove, infrastrukturu i BD arhitekturu, istraživanje mogućnosti uvođenja novih tehnologija za eksploataciju, ali i primjenu dobijenih outputa s ciljem unapređenja različitih poslovnih procesa, uzimajući u obzir specifičnosti različitih djelatnosti.

IZVORI

1. Aleksić-Marić, V., 2008. *Poslovna informatika*. Banja Luka: Ekonomski fakultet, Banja Luka.
2. Apache, 2015. Hadoop: Apache. [Na mreži] Dostupno na: <http://hadoop.apache.org/>
3. Appelbe, B., 2014. *Hadoop & Big Data: Learning Tree*. [Na mreži] Dostupno na: https://www2.learningtree.com/training/lp/big-data-t.aspx?redirect=1&v_url=WhatIsBigData
4. Bedgood, L., 2014. *Fast Data: The Next Revolution of Big Data*. [Na mreži] Dostupno na: <http://www.datamentors.com/blog/fast-data-next-revolution-big-data>
5. Cisco, 2013. Cisco Global Cloud Index: Forecast and Methodology 2013–2018 White Paper. [Na mreži] Dostupno na: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html
6. Empolis, 2015. Big Data and Extreme Information Management. [Na mreži] Dostupno na: <http://www.empolis.com/en/smart-information-management/technology/big-data-and-extreme-information-management.html>
7. Harvey, C., 2012. *Data Center: Datamation*. [Na mreži] Dostupno na: <http://www.datamation.com/data-center/50-top-open-source-tools-for-big-data-1.html>
8. IBM Big Data Hub, 2015. Infographics & Animations, The Four V's of Big Data. [Na mreži] Dostupno na: <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>
9. IBM, 2014. Big Data: IBM. [Na mreži] Dostupno na: <http://www.ibm.com/big-data/us/en/technology/>
10. IBM, 2014. Xtreme Data. [Na mreži] Dostupno na: <https://www.xtremedata.com/xtremedata-apr-14>
11. Intel, 2013. Learn Big Data and Make a Big Impact. [Na mreži] Dostupno na: <http://www.intel.co.uk/content/www/uk/en/big-data/learn-about-big-data.html>

12. Mariner , 2014. Why “Big Data” is Causing Enterprise Impotence and How to Cure it. [Na mreži] Dostupno na: <http://www.mariner-usa.com/blog/big-data-causing-enterprise-impotence-cure/>
13. McKinsey Global Institute, 2011. Big data: The next frontier for innovation, competition, and productivity, s.l.: s.n.
14. Miller, M. J., 2011. *Beyond Big Data: Gartner on “Extreme” Information Management*. [Na mreži] Dostupno na: <http://forwardthinking.pcmag.com/show-reports/289401-beyond-big-data-gartner-on-extreme-information-management>
15. Mir, H., 2013. *The Value Amazon, Facebook, and Google are Extracting From the Big Data*. [Na mreži] Dostupno na: <http://zerotoprotraining.com/index.php?mode=video&id=330>
16. Mir, H., 2013. *Tutorials: Zero To Pro Training*. [Na mreži] Dostupno na: <http://zerotoprotraining.com/index.php?mode=video&id=327>
17. Mir, H., 2014. *Big Data Risks*. [Na mreži] Dostupno na: <http://zerotoprotraining.com/index.php?mode=video&id=1181>
18. Nelsen, R., 2015. *Big Data is Old News: Enter Fast Data*. [Na mreži] Dostupno na: <https://www.qualtrics.com/blog/big-data-is-old-news-enter-fast-data/>
19. Pilcher, M., 2010. *Extreme Data*. [Na mreži] Dostupno na: <http://www.sand.com/extreme-data/>
20. Smith, T., 2013. You Tube. [Na mreži] Dostupno na: <https://www.youtube.com/watch?v=j-0cUmUyb-Y>
21. Stats, I. W., 2015. *Stats: Internet World Stats*. [Na mreži] Dostupno na: <http://www.internetworldstats.com/stats.htm>
22. Telefonica, 2014. Press: m2m Telefonica. [Na mreži] Dostupno na: <https://m2m.telefonica.com/press/m2m-800-million-electric-smart-meters-to-be-installed-globally-by-2020>
23. Typesafe, 2015. New Typesafe White Paper Explores Evolution of Big Data to “Fast Data”. [Na mreži] Dostupno na: <https://www.typesafe.com/company/news/new-typesafe-white-paper-explores-evolution-of-big-data-to-fast-data>
24. Vaughan-Nichols, S. J., 2015. *The most popular web browsers: ZD Net*. [Na mreži] Dostupno na: <http://www.zdnet.com/article/the-most-u-s-popular-web-browsers/>
25. Weinman, J., 2013, *Cloudonomics: The Business Value of Cloud Computing*. John Wiley and Sons. Kindle Edition.