

Big Data i poslovna inteligencija

Rastko Pavlović
m:tel a.d.
Banja Luka
rastko.pavlovic@mtel.ba

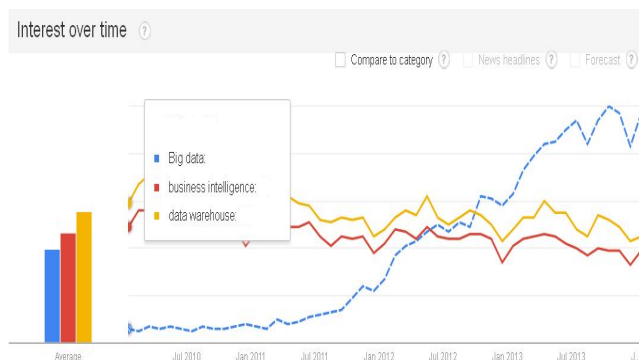
Ratko Dejanović
Elektrotehnički fakultet
Banja Luka
ratko@etfbl.org

Sadržaj— Big Data je termin iz informacionih tehnologija, tačnije poslovne analitike kao njihovog dijela, za koji je u nekoliko zadnjih godina po Google Trends-u pokazan najveći interes. U ovom radu se analizira Big Data koncept sa posebnim osvrtom na mogućnosti njegove primjene u sistemima poslovne inteligencije.

Ključne riječi— Big Data, Poslovna inteligencija; skladištenje podataka

I. UVOD

Zadnjih nekoliko godina u informacionim tehnologijama su više nego dinamične, aplikacije i servisi se pomjeraju u oblak, obrada transakcionih podataka se proširuje i obradom interaktivnih podataka sa socijalnih mreža kao što su Facebook, Twitter, LinkedIn, a uz sve to ide promjena “ulaznog kanala” do aplikacija i informacija sa desktop računara na mobilne komunikatore. Razni oblici mobilnih komunikacionih uređaja se sve više koriste i za pristup aplikacijama, a ne samo za ostvarivanje komunikacije. Računarska paradigma se pomjerila i na mobilne uređaje gdje sa Android operativnim sistemom svako može da “sklopi” mobilni uređaj kao što su sklapani personalni računari, a sa druge strane i da razvije aplikaciju, te sada imamo eksploziju aplikacija i informacija. U posljednje tri godine ljudsko društvo je stvorilo i sačuvalo više informacija nego u čitavoj istoriji do tada [1].



Slika 1. Interesovanje za termin Big Data [1]

Ono što se ne može promijeniti, već samo može da se koristi u značajnijoj mjeri je potreba da se donošenje odluka zasniva na prikupljenom znanju, ali sam način prikupljanja i obrade informacija neophodnih za odlučivanje mora da se prilagodi novom okruženju.

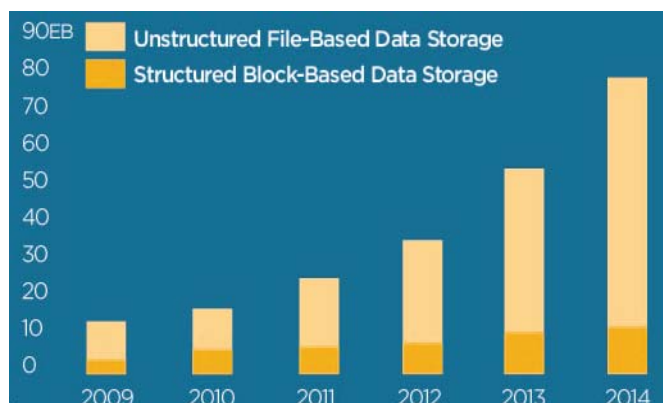
U ovom radu se analizira Big Data koncept i njegov uticaj, odnosno iskoristivost ovog koncepta u postojećim sistemima poslovne inteligencije za donošenje poslovnih odluka.

II. BIG DATA KONCEPT

Opseg onoga što se danas posmatra kao Big Data je širok, a definicije su nejasne čak i kontradiktorne. Najšire prihvaćena definicija termina Big Data proistekla je iz analize Meta Grupe (sadašnji Gartner) koja je rađena 2001 godine i po toj definiciji pod pojmom Big Data se podrazumijeva informacioni resurs velike količine, velike brzine i velike raznovrsnosti podataka koji zahtijeva nove i inovativne metode obrade i optimizacije informacija, poboljšanje uvida u sadržaj podataka i donošenja odluka [3]. Ovo je definicija tzv. 3V dimenzije Volume, Variety, Velocity (količina, raznovrsnost, brzina) gdje količina podataka ide i do reda peta bajtova, a raznovrsnost podataka koji se obrađuju u realnom vremenu je široka (strukturirani, nestrukturirani-tekst, multimedija). Neki autori definiciju Big Data proširuju na još jednu dimenziju 4V veracity - istinitost (pouzdanost, predvidljivost, tačnost podataka), a neki i na 6V gdje dodaju dimenzije viscosity – djelotvornost i value – vrijednost. Definicija koju je potrebno izdvojiti, koja je pokušaj da se objedini više definicija - “Big Data je pojam koji označava smještanje i analizu velike, kompleksne strukture podataka korištenjem niza tehnika, uključujući, ali ne ograničavajući se na tehnologije NoSQL, MapReduce i mašinsko učenje” [4]. Iza definicija se ne kriju samo problemi obrade veće količine raspoloživih podataka, kao što sam termin Big Data može da asocira, koji su nastali kao rezultat promjena u informacionim tehnologijama, već podrazumijevaju i unapređenje analitičkih kapaciteta i analiza svih podataka koji bi se koristili u poslovnim procesima kao i procesu odlučivanja. Kakva god definicija bila, glavni pokretač i razlog za unapređenje ove tehnologije, da li ono bilo tu ili ne, je V koje označava value-vrijednost koju ona donosi. Prepoznavanje vrijednosti podataka- “Data as a asset” – odnosno podataka kao resursa sa kojima se raspolaže nije došao sa ovim konceptom, ali ovaj koncept može otvoriti novu zavjesu iza koje leže veliki dodatni resursi koji se mogu iskoristiti. Značaj podataka, odnosno Big Data koncepta nije prepoznat samo u poslovnom svijetu, vlade mnogih zemalja pokušavaju da se pozicioniraju i iskoriste prednosti posjedovanja podataka te su u 2013 godini donijeli strategije u ovoj oblasti. Tako je australijska vlada usvojila Big Data strategiju u kojoj je definisila da su podaci koje oni posjeduju resurs od nacionalnog značaja “Data is a National asset” [5], a slične dokumente su donijele i druge zemlje (UK, USA, Japan) [6]. Po istraživanju samo 7% ispitanika je u potpunosti zadovoljno sa stepenom

obuhvaćenosti podataka koji se koriste u postojećim sistemima i koji ulaze u dalje analize i na osnovu kojih se donose odluke [7]. Big Data koncept za analize podataka koje se poduzimaju podrazumijeva veliku količinu podataka iz različitih izvora, uključujući struktuirano, polu-struktuirani, nestruktuirani ili čak i nepotpune podatke, a veličina podataka u okviru analize i brzina koja je potrebna kako bi se analizirala prevazilazi sposobnosti standardnih alata i metoda analize. Ono što je još jednom potrebno istaći, osim veličine i raznovrsnosti, koje kako je rečeno karakterišu Big Data koncept, je brzina obrade podataka koja se pomjera u realno vrijeme i po ovom konceptu informacija je potrebna u trenutku kada se događaj odigrava.

Po nekim procjenama količina podataka koji se kreiraju će biti 44 puta veća u 2020 nego što je bila u 2009 [8], te samim tim način na koji se oni prikupljaju i obrađuju se mijenja, a ako se na to doda i sve veći udio nestruktuiranih podataka onda je jasno da se mora na drugačiji način pristupiti ovim podacima. Paradigma Big Data je bila glavna tema u 2012 godini, dok se u 2013 godini rasprava pomjerila na temu



Slika 2. Procjena odnosa sačuvanih struktuiranih i nestruktuiranih podataka [9]

vrijednosti znanja koji se nalazi u tim podacima i tu se otvaraju mnoga pitanja: kako kontrolisati troškove eksponencijalnog rasta količine podataka, promjene tipove podataka, načina korištenja istorijskih i podataka u realnom vremenu, jednom riječju vrijednosti koje može donijeti ovaj koncept. Ono što je činjenica je da on postoji, i da su čak i države donijele strategije vezane za Big Data, ali ostaje otvoreno pitanje ko da ga koristi i kako da ga iskoristi? Dva su mjesta gdje se on koristi, operacioni sistemi i sistemi poslovne inteligencije, ali bilo gdje da je, samo ako donosi dodatnu vrijednost može da živi i da se dalje razvija.

III. BIG DATA TEHNOLOGIJA

Sami početak primjene od prije desetak godina vezuje se za kompanije kao što su Google, Yahoo koje su imale ogromne količine podatak generisanih na webu, da bi 2008 godine Hadoop postao Apache projekat otvorenog koda najvišeg nivoa što je dalo presudan momentum za razvoj ove tehnologije. Hadoop i MapReduce tehnologija su osmišljeni kako bi se

efikasno mogle obraditi velike količine podataka korištenjem više računara vezanih u klaster. MapReduce se zasniva na okruženju u kojem se protok podataka razbija u dvije faze, fazu mapiranja i fazu redukcije. U fazi mapiranja, grupe podataka se obrađuju izolovano posebnim procesima koji se označavaju kao maperi gdje se pojedinačni elementi dijele u uređene parove tj. ključ-vrijednost parove. Rezultat rada mapera se prosleđuje procesima redukcije, koji kombinuju ove uređene parove u manje skupove što su i konačni rezultati. Na strani smještanja podataka podršku za ovu tehnologiju klasterisanja podataka na više sistema pruža Hadoop distribuirani fajl sistem (HDFS). Podaci u Hadoop klasteru se dijele u manje dijelove (tzv. blokove) i distribuiraju se kroz klaster, te se na taj način omogućava da se procesi mapiranja i redukcije rade na manjim količinama podataka čime se dobija na skalabilnosti. Sama ideja je bila da se koriste serveri u klasteru sa jeftinim internim diskovima što može da predstavlja problem zbog brzine i dostupnosti, odnosno gubitka podataka. Cilj je da se MapReduce poslovi izvršavaju na serverima gdje se nalaze podaci koji se obrađuju tzv. princip lokalnih podataka. Kako HDFS prilikom smještanja podatke dijeli na manje blokove i kopije ovih blokova se smještaju i repliciraju se na različite servere u okviru klastera (minimalno dva servera) dobija se veća dostupnost podataka i obrada na svim serverima. Cjelokupna logika upravljanja podacima se obavlja na serveru koji se označava kao NameNode. Ovaj server je kao neka vrsta servera sa meta podacima u kojima se drže informacije o fajlovima, odnosno blokovima podataka na Hadoop distribuiranom fajl sistemu. Sve informacije koje ima NameNode server se drže u memoriji što omogućava brzu manipulaciju sa podacima koji su na fajl sistemu ili se upisuju na njega, ali sa druge strane predstavlja kritičnu tačku prekida, jer bi se gubitkom meta podataka izgubile sve informacije o podacima u klasteru. Osim što ovaj server treba biti robusniji na prekide od servera u klasteru potrebno je obezbijediti i njegov bekap. Od verzije Hadoop 0.22 je implementirana Backup Node funkcionalnost koja obezbjeđuje održavanje ažurne kopije NameNode u memoriji. Pet elemenata čini Big Data tehnologiju analitička baza podataka, memorijska baza (In-memory database management system), obrada tokova podataka u realnom vremenu, Hadoop te softver i hardver na kojem se to sve smješta.

Kao podrška za ovu tehnologiju u okviru Apache projekta se razvija više projekata otvorenog koda povezanih sa Hadoop-om: Ambari, Avro, HBase, Cassandra, Chukwa, Mahout, Hive, Pig, Spark. Ambari je Apache projekat koji pruža podršku za jednostavnije upravljanje (kreiranje, pokretanje, nadgledanje, konfigurisanje) Hadoop klasterom kroz web baziran interfejs. Avro je sistem za serijalizaciju podataka. Cassandra je Hadoop distribuirana baza podataka (database system) otvorenog koda koji koriste Ebay, Twwitter i mnogo drugih kompanija čija je karakteristika da imaju velike količine aktivnih podataka. Najveći poznati Cassandra klaster ima preko 300 TB podataka na više od 400 servera. HBase je takođe Hadoop distribuirana, skalabilna baza za smještanje velikih količina podataka. HBase i Cassandra su vrsta "NoSQL" baza podataka koje su distribuirane.

"NoSQL" je pojam koji označava da baza podataka nije relacionala baza koja podržava SQL (structured query language) kao svoj primarni jezik. Chukwa je sistem prikupljanja podataka za praćenje velikih distribuiranih sistema. Mahout projekat ima cilj izgradnje skalabilne biblioteke mašinskog učenja. Hive omogućava upite i upravljanje velikim skupovima podataka koji se nalaze u distribuiranom sistemu. Pig je, zajedno sa infrastrukturom koja to podržava, platforma za analize velikih količina podataka. Spark je okruženje za brzu obradu podataka, dok Impala, okruženje za paralelnu obradu podataka zaobilazi MapReduce kako bi se obezbijedila mogućnost upita u realnom vremenu, a samim tim omogućava i analitiku, odnosno pristup alatima poslovne inteligencije nad podacima smještenim u Hadoop-u u realnom vremenu. Iznad svih njih se nalazi ZooKeeper, centralizovani servis za upravljanje i konfigurisanje ove Apache tzv. "ZOO" infrastrukture.

IV. BIG DATA KAO DIO SISTEMA POSLOVNE INTELIGENCIJE

Skladišta podataka kao gradivni dio sistema poslovne inteligencije su tako organizovane baze podataka (database systems) da kroz alate poslovne inteligencije omogućavaju uvid u poslovanje i donošenje poslovnih odluka zasnovanih na činjenicama i oslobođenih subjektivnih impresija. Kako bi se ovi sistemi adaptirali na promjene koje donosi Big Data, u kontekstu donošenja poslovnih odluka, jasno je da je neophodna neka vrsta nadogradnje i širenje postojećih sistema kako bi se obuhvatili novi izvori sa novim tipovima podataka i značajno veće količine podataka koje nisu striktno struktuirane. Ove promjene trebaju da pomjere granice poslovanja i na njega ne samo da dodaju novu vrijednost, već i odluka koja je proizvodi i informacije na kojima se zasniva trebaju da budu u realnom vremenu. Kada korisnik ostvari kontakt putem kontakta centra ili dođe na prodajno mjesto, da bi se na najbolji način iskoristio već ostvareni kontakt i korisniku dala nova prodajna ponuda, osim istorijskih podataka koji se nalaze u skladištu podataka može se iskoristiti i trenutna informacija i sklopiti mu više verzija ponuda koje su najbolje za prodavca. Ponuda koja je uobzirila i prethodno ponašanje korisnika ima veliku vjerovatnoća da će je korisnik prihvatiti. Zavisno od poslovnog modela i mjesta gdje se odluke donose potrebno je upariti podatke koji dolaze u realnom vremenu sa istorijskim podacima i isporučiti informacije na način kako poslovanje zahtijeva. Po izvedenom ispitivanju 13% kompanija koriste prediktivnu analizu, 75% postojećih skladišta podataka ne mogu da podrže brzinu i kompleksnost, a 86% kompanija ne mogu da isporuče pravu vrijednost u pravo vrijeme [10]. Ove nedostajuće procenta u značajnijoj mjeri moguće je "popuniti" sa Big Data konceptom. Sistemi poslovne inteligencije su u stalnom procesu evolucije u kojem se u njih integrišu različita tehnološka unapređenja tako da je proširenje na Big Data koncept, prirodan razvoj sistema poslovne inteligencije. Dva su mjesta u sistemu poslovne inteligencije gdje Big Data koncept može donijeti poboljšanja, Hadoop i vertikalne baze

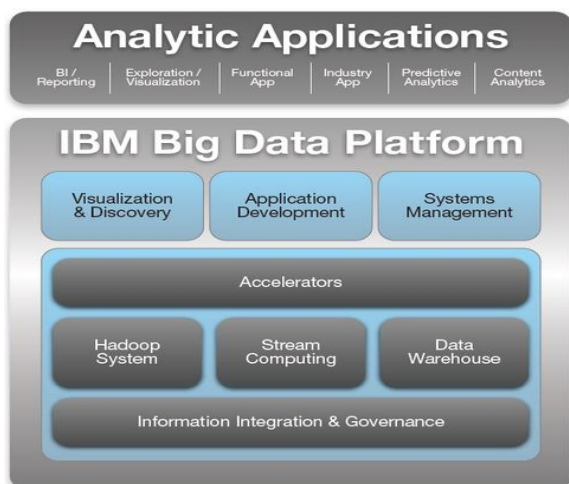
podataka u pripremnoj oblasti tzv. Staging area, te na prezentacionom nivou uz memorijske baze (in-memory) gdje su vizualizacija i analitičke aplikacije prilagođene krajnjem korisniku.

Arhitektura zasnovana na Apache Hadoop projektu otvorenog koda koja omogućava da se obrade velike količine podataka na klasterima servera, čak je bila prepoznata i kao osnova sledeće generacija sistema poslovne inteligencije, mada je u praksi, u ovim sistemima, za sada najčešće iskorištena njegoova funkcionalnost da obrađuje velike količine nestruktuiranih podataka za punjenje postojećih skladišta kao neka vrsta ETL (ekstrakcija, transformacija, punjenje) procesa. Ova tehnologija kroz skalabilnost smještanja podataka i brzine obrade podataka omogućava analize brzinama i cijenama koje prije nisu mogle biti postignute što je obesmišljavalo njihovu obradu. Još jedna značajna prednost je što je tehnologija otvorenog koda, koja može da smanji troškove licenciranja i održavanja. Kako je uvijek bilo teško jednoznačno procijeniti povrat ulaganja (ROI-return of investment) kod projekata sisteme poslovne inteligencije, tako je i sa projektima njihovog proširenja sa Big Data konceptom. Da li je to uzrok što se ovi projekti ne pokreću sa brzinom koja bi bila u neku ruku proporcionalna količinama informacija, ili je sa druge strane količina informacija i priča koje su vezane za Big Data u zadnje vrijeme toliki da se stiče utisak da je ovaj koncept bio opšteprihvaćen mnogo ranije, te kratko vrijeme za koje je on prisutan ne može biti razmatran kao razlog za to. Veliki podaci su postojali, ali tek sa tehnološkim preduslovima i prepoznavanjem vrijednosti koja se sistemima poslovne inteligencije donosi od velikih podataka postao je Big Data koncept.

V. BIG DATA RJEŠENJA

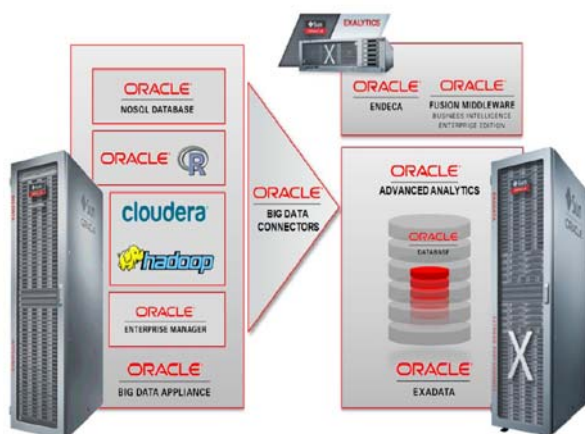
Tehnologija Big Data i prednosti koje ona donosi prepoznata je od vodećih softverskih kompanija koje isporučuju komercijalni softver. Talas kojeg nosi Big Data pokušavaju da iskoriste, te su dali svoje viđenje ovog koncepta i nadogradnju na Hadoop okruženje i iskoristili su ovu ideju kako bi proširili svoje poslovanje, odnosno isporučili nova rješenja. U zavisnosti od njihove ponude softvera zavisilo je i mjesto i način implementacije Big Data koncepta.

IBM je u oktobru 2010 objavio svoju IBM InfoSphere BigInsights platformu sa ciljem da Hadoop bude dostupan njihovim poslovnim softverskim rješenjima sa tri glavna aspekta analitike, integracije podataka i upravljanja ovim sistemima. Platforma je postavljena na način da analitika velikih podataka koezistira sa analitikom na drugim vrstama podataka, a Hadoop klasteri nesmetano rade svoj posao paralelno sa IBM serverima. BigInsights osigurava interoperabilnost s DB2, Netezza, IBM Smart Analytics sistemima skladištenja podataka. IBM rješenje omogućava čvrstu integraciju gdje se upiti mogu automatski distribuirati između DB2 i Hadoop po potrebi.



Slika 3. IBM Big Data platform [11]

Osim korištenja open source softvera, kao što su Pig, Hive, HBase, Avro, Lucene i Oozie, IBM je razvio i nove alate kao što su BigSheets, Jaql i SystemT napredne analitike koji omogućavaju upravljanje u Hadoop okruženju. [1] Oracle Big Data strategija je slična i zasniva se na ideji proširenja postojeće informacione arhitekturu kako bi se omogućila podrška za Big Data koncept. Nove Big Data tehnologije, kao što su Hadoop i Oracle NoSQL baze podataka, rade uz postojeće Oracle skladište podataka kako bi isporučili vrijednost i znanje i zadovoljile zahtjeve za podacima. Kao NoSQL baza podataka koristi se distribuirana, skalabilna Oracle Berkeley. Oracle je omogućio Big Data dodatke na način da oni omogućavaju prikupljanje i organizaciju novih tipova podataka, Oracle Big Data konektori su ti koji integrišu Big Data okruženje sa



Slika 4. Oracle Big Data platform [12]

njihovom Oracle Exadata i Oracle bazom podataka. Oracle Loader for Hadoop (OLH) omogućava korištenje Hadoop MapReduce obrade kako bi se kreirali optimizovani podaci za efiksano punjenje i analizu podataka u Oracle bazi podataka. Za razliku od drugih sistema za pristup podacima u Hadoop

okruženju, ovdje se generiše Oracle-ov interni format kako bi se omogućilo brže učitavanje podataka sa što manjim korištenjem resursa baze podataka. Za direktni pristup podacima velikim brzinama koji se nalaze na Hadoop distribuiranom fajl sistemu direktno iz Oracle baze podataka koristi se Oracle SQL Connector for Hadoop. Kada su podaci učitani u Oracle bazu podataka korisnicima su na raspolaganju svi dodatni Oracle alati za analitiku označeni kao in-database (In-Database Data Mining, Oracle R Enterprise, In-Database Text Mining, In-Database MapReduce), a kroz Oracle Exalytics In-Memory Machine obezbijedene su visoke performanse analitike u memoriji [12].

SAS kao sveobuhvatna platforma poslovne inteligencije sa svim svojim komponentama koje imaju za cilj povećanje poslovnih performansi svoj Big Data koncept razvoja temelji na sveobuhvatnosti čitavog životnog ciklusa podataka koristi High-Performance Analytics za Hadoop podršku

Za integraciju i transformaciju podataka SAS sistema i Hadoop-a koristi se SAS Data Federation Server, dok SAS Event Stream Processing Engine podržava kompleksne obrade velike količine podataka u realnom vremenu. Kao što se i podrazumjeva za sisteme poslovne inteligencije na vrhu je analitički alat SAS Visual Analytics koji kroz pristup Hadoop klasteru i memorijskoj bazi (in-memory) omogućava vizualizaciju i analize podataka. SAS koncept Big Data je na području Bosne i Hercegovine implementiran u UNICREDIT banci, a u toku je implementacija u m:tel—u.

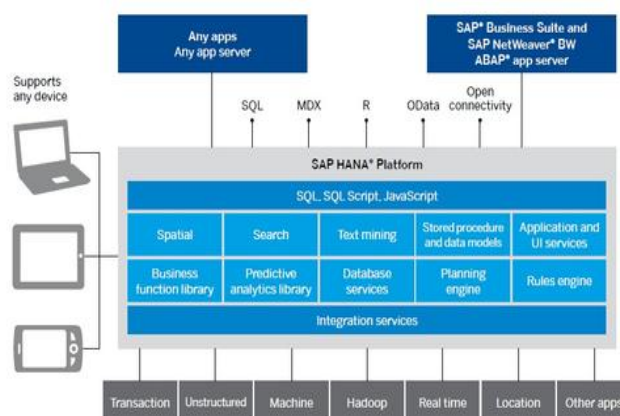


Figure: SAP HANA – An in-memory analytics and applications platform for real-time business

Slika 5. SAP Big Data platforma **Error! Reference source not found.**

SAP je jedan od lidera u ovoj oblasti i razvili su obiman skup rješenja za Big Data koncept koji omogućava integraciju Big Data koncepta u postojeće sisteme poslovne inteligencije i skladišta podataka na više načina. Centralna tačka njihovog rješenja je integracija Hadoop-a zajedno sa SAP HANA, memorijskom bazom koja omogućava analitiku u realnom vremenu, a obradu podataka u realnom vremenu omogućava SAP Sybase Stream processor. Pristup podacima koji se nalaze na Hadoop distribuiranom fajl sistemu koji se koriste da li za SAP HANA ili Sybase IQ (vertikalna baza podataka) omogućen

je kroz SAP Data Services. SAP HANA kroz svoju funkcionalnost „smart data access“ omogućava i postavljanje direktnih upita na Hadoop Hive. a alati poslovne inteligencije SAP BusinessObjects , SAP Visual Intelligence i SAP Predictive Analysis takođe mogu direktno da pristupe Hadoop okruženju [13]. U testnom okruženju SAP je u 2013 godini postigao brzinu učitavanja i indeksiranja od 34.3 terabajta podataka za sat vremena **Error! Reference source not found.**

Prednosti ovih i drugih komercijalnih rješenja su ta da korisnicima koji imaju znanje za rad u postojećim okruženjima, IBM, Oracle, omogućiti i rad sa Big Data

VI. ZAKLJUČAK

Big Data koncept možemo posmatrati kao novi element koja se može ugraditi u različite tipove informacionih sistema. Da li ćete sa njim rješavati problem operativnog poslovanja i koristiti njegove funkcionalnosti za operativne poslove kao što je na primjer detekcija prevara (fraud detection) ili poslovnu inteligenciju, ili i jedno i drugo, zavisi od strategije razvoja informacionog sistema. Ako ga posmatramo u kontekstu donošenja odluka, Big Data je dodatni gradivni element u sistemu poslovne inteligencije koji potencijalno može donijeti novo znanje, a sa njim i novu vrijednost. Poslovna inteligencija je koncept koji unapređuje mehanizam donošenja poslovnih odluka, a to na koji način će se prikupljati i obrađivati , da li je urađen ovakav ili onakav ETL (ekstrakcija, transformacija, punjenje) proces nije presudno. Analogno skladištu podataka koje nije sistem poslovne inteligencije već njegov sastavni dio tako je i Big Data, kao i bilo šta drugo što unapređuje informisanost, postalo dio sistema poslovne inteligencije.

Kao što dijelovi sistema skladištenja podataka ne moraju samo da se koriste u sistemu poslovne inteligencije, isto tako i Big Data koncept nije nužno vezan za poslovnu inteligenciju , ali ako pričamo o donošenju poslovnih odluka onda mu je tu mjesto.

Big Data ne može da bude zamjena za sisteme poslovne inteligencije i skladište podataka, ali može kroz postojeće sisteme da donese novu vrijednost. Koliko će se znanja izvući i koja je to mjera da se negdje primjeni ili ne Big Data koncept još je u domenu teorije pošto je praksa na svojim počecima .

Big Data koncept u poslovnoj inteligenciji se nalazi u fazi u kojoj su ispirajući zlata bili na počecima zlatne groznice, koliko pijeska i na kojim mjestima treba da se obradi kako bi se našao grumen zlata zavisi od mnogo faktora. Više nije pitanje opravdanosti ovog koncepta i njegove iskoristivosti u poslovnoj inteligenciji, zavjesa je skinuta, a koji su to faktori koji definišu gdje, šta i kako da se obradi će doći kroz korištenje i iskustvo koje iz njega proistekne.

LITERATURA

- [1] Paul C. Zikopoulos, Dirk deRoos, Krishnan Parasuraman, Thomas Deutsch, David Corrigan, James Giles, „Harness the Power of Big Data“ , McGrawHill, 2012
- [2] Google Trends
- [3] Doug Laney, „Application Delivery Strategies“, Meta Group, 2001
- [4] Jonathan Stuart Ward and Adam Barker, „Undefined By Data: A Survey of Big Data Definitions“ School of Computer Science University of St Andrews, UK, September 2013
- [5] „The Australian Public Service Big Data Strategy“, Department of Finance and Deregulation, August 2013
- [6] „A strategy for UK data capability“ , HM Government, October 2013
- [7] „Next generation customer experience management“, Aberdeen group, Mart 2013
- [8] Bruno Aziza, „Predictions for Big Data“, April 2013
- [9] „Data rEvolution“, CSC LEADING EDGE FORUM, September 2011
- [10] „Worldwide File-Based Storage research“, IDC, 2013
- [11] <http://www-01.ibm.com/software/data/bigdata/>
- [12] „Big Data for the Enterprise“, ORACLE, June 2013
- [13] <http://scn.sap.com>
- [14] Philip Russom : „Managing Big Data“, TDWI, September 2013

ABSTRACT

Big Data is a term from business analytics (part of information technology), for which is shown the greatest interest in the last few years (Google Trends). This paper analyzes the concept of Big Data with special reference to the possibility of its application in business intelligence systems.

BIG DATA AND BUSINESS INTELLIGENCE

Rastko Pavlović
Ratko Dejanović