

**ARHITEKTURA SISTEMA POSLOVNE INTELIGENCIJE
BUSINESS INTELLIGENCE SYSTEM ARCHITECTURE**Bojana Popov, *Fakultet tehničkih nauka, Novi Sad***Oblast – INDUSTRIJSKO INŽENJERSTVO I
INŽENJERSKI MENADŽMENT**

Kratak sadržaj – Poslovna inteligencija je naziv za proces prikupljanja, obrade i čuvanja podataka koji je podržan računarskim tehnologijama. Poboljšava proces donošenja odluka pružajući menadžmentu relevantne informacije o poslovanju. U radu je prikazana arhitektura sistema poslovne inteligencije i koristi nakon implementacije.

Ključne reči: Poslovna inteligencija, BI architecture, skladište podataka, analiza podataka

Abstract Business intelligence is the name for a process of collecting, processing and storing data that is supported by computer technologies. Improves the decision-making process by providing management with relevant business information. The paper presents the architecture of the business intelligence system and its benefits after implementation.

Key words: Business intelligence, BI architecture, data warehouse, data analysis

1. UVOD

Proces donošenja odluka u modernim organizacijama u velikoj meri se oslanja na informacione sisteme. Podaci koji se generišu iz internih i eksternih izvora obrađuju se kako bi se dobile vredne informacije koje utiču na donosiocima odluka. Prvobitno, informacioni sistemi kakav je MIS (eng. *Management Information System*) ili upravljački informacioni sistem služio je za skladištenje svakodnevnih operacija. Ovaj sistem nije čuvao i skladištio istorijske i agregirane podatke. Prema tome, izveštaji koji su se generisali zahvaljujući ovom sistemu bili su veoma jednostavni i generični. Pružao je skromne mogućnosti pri obradi podataka – služio je za analizu operativnih aktivnosti. Sistemi za podršku odlučivanju (eng. *Decision Support Systems*) nastali su kako bi se rešili problemi u vezi sa upravljačkim informacionim sistemima i kako bi se omogućila analiza podataka na višem nivou. Ovaj sistem ima jedinstveno skladište podataka u kojem su ujedinjeni svi podaci organizaciji, nasuprot MIS-u, gde svaka poslovna funkcija zasebno čuva podatke o svom poslovanju. Zahvaljujući centralizovanom skladištenju podataka, unapređene su i mogućnosti analize.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Zdravko Tešić, red. prof.

Sistemi za podršku odlučivanju sadrže alate za analitičku obradu na mreži. Ovi alati smeštaju podatke u specijalizovanu formu koja se naziva OLAP (*Online Analytical Processing*) kocka i pružaju multidimenzionalni pogled na podatke. Uz pomoć OLAP tehnologija i tehnika rudarenja podataka moguće je pretvoriti ulazne podatke u znanje i informacije na osnovu kojih se donose odluke. Koncepti sistema za podršku odlučivanju i poslovna inteligencija se često u literaturi koriste zajedno.

Neki autori smatraju da je sistem poslovne inteligencije naslednik sistema za podršku odlučivanju. Drugi međutim, smatraju da je sistem za podršku odlučivanju jedan od elemenata poslovne inteligencije. Poslovna inteligencija ima sve komponente koje sadrži DSS, a pored toga sadrži bogat korisnički interfejs koji omogućuje da obrađeni podaci budu prikazani putem različitih kontrolnih tabli. Kontrolne table daju višeperspektivni pogled na skup podataka sa vizuelizacijama koje predstavljaju različite nalaze i uvide iz tog skupa podataka. Vizuelizacije se mogu zakačiti na kontrolne table i ako se odabere zakačena vizuelizacija, otvoriće se izveštaj vezan za nju.

**2. OPŠTA ARHITEKTURA SISTEMA POSLOVNE
INTELIGENCIJE**

Postoje različite definicije koje opisuju poslovnu inteligenciju, a termin je popularizovao Howard Dresner 1989. godine i on glasi „BI opisuje skup koncepata i metoda za poboljšanje poslovnog odlučivanja korišćenjem sistema podrške zasnovanih na činjenicama. Sistemi poslovne inteligencije su DSS vođeni podacima“ [1]. Pregledom literature utvrđeno je da autori ne mogu da se slože oko jedinstvene strukture komponenti BI (*Business intelligence*) sistema [2][3][4][5]. Međutim, iako postoje neka neslaganja, većina autora se slaže da se BI sastoji iz sledećih slojeva:

- sloj podataka,
- ETL sloj,
- sloj skladišta podataka,
- sloj grafičkog korisničkog interfejsa i
- sloj metapodataka.

BI arhitektura se može primeniti u lokalnom centru podataka ili na oblaku (eng. *cloud*). U oba slučaja sadrži skup navedenih komponenti koje podržavaju različite faze procesa. Proces započinje prikupljanjem podataka iz različitih izvora, zatim se ti podaci prečišćavaju i integrišu kako bi mogli da se učitaju u skladište podataka sa definisanim šemom. Iz skladišta podataka se dalje izvlače kako bi se nad njima primenile određene tehnike analize. Na kraju, rezultati koji su dobijeni analizom prikazuju se u vidu izveštaja ili vizuelizacija.

2.1. Sloj podataka

Sloj podataka čine podaci koji se generišu iz različitih izvora. Izvori podataka ne čine deo platforme poslovne inteligencije. Oni reprezentuju podatke koji se dobijaju iz internih i eksternih izvora. Interne izvore čine svi podaci koji se generišu u okviru same organizacije. To mogu biti operativne baze podataka, sistemi za upravljanje odnosima sa klijentima (eng. *Customer Relationship Managment*), sistemi za planiranje resursa preduzeća (eng. *Enterprise Resource Planning*), finansijski izveštaji i drugi interni dokumenti. Spoljni izvori se odnose na podatke koji potiču van organizacije. To su podaci koji se prikupljaju od strane kupaca, poslovnih partnera, vlade, organizacija za istraživanje tržišta i sa Interneta.

Svi podaci mogu se klasifikovati u tri kategorije: strukturirani podaci, polustrukturirani podaci i nestrukturirani podaci. Strukturirani podaci imaju definisanu dužinu i format. Strukturirani podatak o nekom entitetu predstavlja jedan red u relacionoj bazi podataka. Nestrukturirani podaci se ne mogu skladištiti u tradicionalnoj relacionoj bazi. Zbog toga, ove podatke je teže analizirati i do skoro nisu imali veliku vrednost kod organizacija.

One organizacije koje uspešno koriste nestrukturirane podatke kako bi donosile poslovne odluke, stiču značajnu prednost na tržištu. Najčešći formati nestrukturiranih podataka su pdf, video, audio i veb sadržaj. Polustrukturirani podaci su hibrid strukturiranih i nestrukturiranih podataka. To je oblik podataka koji se ne može skladištiti u tradicionalnim, relacionim bazama podataka ili u drugim oblicima tabela podataka, ali ipak sadrži oznake za odvajanje semantičkih elemenata i sprovođenje hijerarhije zapisa i polja u okviru podataka. Primeri ovakvih podataka su XML (*Extensible Markup Language*), email i EDI (*Electronic Data Interchange*).

U zavisnosti od tehnologije koja se primenjuje, mora se pažljivo odabrati koji tip podataka može da se koristi u procesu skladištenja, obrade podataka i donošenja odluka. Pored toga, važan kriterijum u procesu odabira podataka podrazumeva relevantnost i kvalitet podataka. Relevantnost podataka podrazumeva sposobnost organizacije da jasno identifikuje izvore podataka koji će biti korisni u procesu rešavanja konkretnog problema. Saznanje gde se traženi podaci, koji su odgovor na specifično poslovno pitanje, mogu pronaći omogućava značajnu uštedu vremena i bržu isporuku informacija. Relevantni izvori mogu biti od pomoći i prilikom obezbeđivanja kvaliteta podataka. Od velikog broja izvora, neki od njih mogu biti nepouzdan, nedostupni ili irelevantni za konkretan problem.

Pravilnim odabirom mogu se izbeći nekonzistentni, netačni, duplirani i podaci koji imaju nedostajuće vrednosti [6]. Međutim, bez obzira na kvalitet podataka koji se nalaze u izvorima, oni se ne učitavaju direktno u skladišta podataka. Podaci se prvobitno iz izvora ekstrahuju u oblast za postavljanje podataka (eng. *Data staging area*) gde se proveravaju kako bi se obezbedio kvalitet i kako bi se formatirali u skladu sa definisanom šemom skladišta podataka u koje će se učitati.

Ova tehnika ekstrakcije podataka, njihovog formatiranja i validiranja i učitavanja u skladište podataka naziva se ETL proces.

2.2. ETL sloj

Kao što je navedeno u akronimu, ovaj sloj se sastoji iz tri faze: ekstrakcije, transformacije i učitavanja podataka. Ekstrakcija je proces identifikovanja i akvizicije relevantnih podataka iz različitih izvora. Prikupljeni podaci često nisu integrisani, nepotpuni su ili su duplirani. U procesu ekstrakcije podaci se iz izvornih sistema učitavaju u oblast za postavljanje podataka kako bi se sprečilo učitavanje netačnih, nepotpunih, dupliranih ili irelevantnih podataka direktno u skladište podataka. Ekstrakcija može biti sinhrona i asinhrona.

Sinhrona ekstrakcija podrazumeva kontinuiranu ekstrakciju, a asinhrona omogućava planiranje ekstrakcije u određenim vremenskim intervalima. Transformacija podrazumeva proces pretvaranja podataka prema definisanim poslovnim pravilima, deduplikaciju i uklanjanje nedostajućih vrednosti podataka. Takođe, oni moraju biti transformisani u konzistentan format za ciljani sistem. Ovo uključuje napore i aktivnosti da se podaci uklupe u višedimenzionalni model koji čini skladište podataka. Zahvaljujući ovom procesu, transformacije mogu biti veoma složene. Poslednja faza je učitavanje podataka u ciljani sistem. Učitavanje podataka može biti inicijalno ili osvežavanje skladišta podataka. Inicijalno punjenje je jednokratna procedura preuzimanja svih istorijskih podataka, a osvežavanje je dodavanje novih podataka koji su se dodali u oblast za postavljanje podatka.

2.3. Sloj skladišta podataka

Sloj skladišta podataka obuhvata tri komponente: skladište operativnih podataka, samo skladište podataka i data mart-ove. Podaci se prenose iz operativnih skladišta u skladište podataka, a zatim u data mart-ove. Operativno skladište podataka (eng. *Operational Data Store*) koristi se za integrisanje svih podataka iz ETL sloja i njihovo učitavanje u skladište podataka. Integriše podatke koji su u realnom vremenu i obezbeđuje pogled na njih. S obzirom da radi sa podacima u realnom vremenu, ne skladišti istorijske podatke. Operativno skladište služi kao podrška operativnoj obradi podataka i za potrebe izveštavanja. Ukoliko određena aplikacija zahteva integrisani pogled na podatke, operativno skladište može ispuniti ovaj zahtev uz pomoć izveštaja [6].

Skladište podataka predstavlja centralizovani repozitorijum gde se smeštaju svi podaci koji se preuzimaju iz izvora podataka. Definicija skladišta podataka koju je konstruisao Bill Inmon glasi: „Skladište podataka je predmetno orijentisano, integrisano, vremenski varijantno, nepromenljivo prikupljanje podataka kao podrška procesu donošenja odluka menadžmenta“ [7]. Predmetna orijentisanost se odnosi na to da se podaci organizuju prema temama poslovanja, nasuprot operativnim bazama podataka, gde su podaci organizovani po funkcijama poslovanja. Ovo je rezultat centralizovanog dizajna skladišta podataka, gde različite funkcije poslovnog sistema pristupaju jedinstvenom skladištu podataka. Na ovaj način jedna tema može biti od interesa za više poslovnih funkcija, a teme mogu deliti zajedničke podatke. Takođe, kada su u pitanju operativne baze podataka, svaka poslovna funkcija može imati različitu implementaciju sistema za upravljanje bazom podataka i podaci mogu biti različito strukturirani. Skladište podataka vodi računa o tome da su podaci u

standardizovanom formatu. Zbog toga skladište podataka ima atribut integrisanosti – svi podaci o organizaciji se čuvaju na jednom mestu, u standardizovanom formatu. Još jedna razlika između operativnih baza podataka i skladišta podataka jeste ta, da skladišta podataka sadrže istorijske podatke. Podaci se dodaju u skladište podataka putem periodičnih osvežavanja – preuzimanja novih podataka iz operativnih baza podataka i ostalih izvora podataka. Dok operativne baze podataka čuvaju podatke do jedne godine, skladišta podataka mogu čuvati podatke do nekoliko godina. Takođe, skladište podataka sadrži vreme kao dimenziju kako bi se podaci vremenski označavali i kako bi se omogućilo kreiranje analiza predviđanja. Nepromenljivost podataka u skladištima podrazumeva da su podaci koji se u njemu nalaze samo za čitanje (eng. *read-only*).

To podrazumeva da u skladištu dozvoljene samo operacije čitanja, ali ne i brisanja ili modifikovanja. Zbog toga, skladišta podataka nekada dostižu velike razmere – do nekoliko petabajta podataka. Ukratko, skladišta podataka sadrže istorijske podatke o celoj organizaciji kako bi omogućili postavljanje upita uz pomoć SQL (eng. *Structured Query Language*) alata i drugih alata za analizu i obradu podataka, koji pomažu menadžmentu u procesu donošenja odluka.

Data mart je deo skladišta podataka koji se fokusira na pojedinačnu poslovnu funkciju, odnosno obuhvata samo jedan segment poslovanja. Koristan je donosiocima odluka jer brže pristupaju podacima i dobijaju neophodan uvid. Korisnici ne moraju da provode vreme pretražujući kompleksnu bazu podataka ili da ručno agregiraju podatke iz različitih izvora. Data mart može biti realizovan kao nezavisni ili zavisni Data mart. Zavisni data mart prikuplja podatke iz centralizovanog skladišta podataka. Nezavisni data mart je dizajniran kao samostalni sistem koji izvlači podatke direktno iz operativnih ili ekstremnih izvora (ili oba).

2.4. Sloj grafičkog korisničkog interfejsa

Na osnovu podataka koji se nalaze u skladištu podataka, u sloju grafičkog korisničkog interfejsa ti podaci se obrađuju i prikazuju donosiocima odluka u vidu izveštaja ili vizuelizacija. Sloj krajnjeg korisnika, kako se drugačije naziva, pruža podršku kako bi se iz podataka dobile smislene informacije i znanje. Za proces obrade podataka i kreiranja informacija koriste se OLAP (*Online Analytical Processing*) kocke i alati za rudarenje podataka. OLAP predstavlja softversku tehnologiju uz pomoć koje korisnici mogu lako i selektivno da izdvajaju i traže podatke kako bi ih analizirali sa različitih tačaka gledišta. Zasnovan je na višedimenzionalnom modelu podataka i omogućava korisnicima da postavljaju upite o višedimenzionalnim podacima.

Podaci se organizuju u tzv. OLAP kocke, a svaka OLAP kocka sadrži podatke kategorisane po dimenzijama (kupci, geografski region, vremenski period) koji su hijerarhijski organizovani. Zahvaljujući ovome, analitičari podataka ne moraju da pišu kompleksne upite kako bi obuhvatili podatke iz više tabela. OLAP kocke takođe agregiraju podatke po dimenzijama, pa je vreme koje je potrebno za odgovor na upite mnogo kraće u poređenju sa kompleksnim upitima koji se postavljaju direktno nad bazom ili skladištem podataka. Osnovne OLAP operacije su:

- Roll-up: penjanje po hijerarhiji koncepata, odnosno povećavanje stepena agregiranosti podataka (ukoliko su podaci prvobitno prikazani po mesecima, uz pomoć roll-up funkcije biće prikazani po godinama).
- Drill-down: spuštanje po hijerarhiji koncepata ili smanjenje stepena agregiranosti podataka (ukoliko su se prvobitno prikazivale zemlje gde organizacija posluje, nakon drill-down operacije, prikazivaće se gradovi u kojima posluje).
- Slice and dice: operacija slice uzima jednu specifičnu dimenziju iz date kocke i predstavlja novu podkocku koja pruža informacije sa nove tačke gledišta. Dice je podkocka koja uzima dve ili više dimenzija.
- Pivot: podrazumeva rotiranje kosa kocke kako bi se prikazale druge dimenzije i informacije koje su vezane za njih.

Rudarenje podataka je računarski podržan metod pomoću kog je moguće otkriti obrasce, pravila, trendove i druge vredne informacije. Rudarenje podataka je proces automatskog pretraživanja obimnih skladišta podataka kako bi se pronašla korisna saznanja. Naziva se još i „otkrivanje znanja podataka“ (eng. *Knowledge Discovery of Data*). OLAP tehnologije omogućavaju otkrivanje veza i obrazaca uz pomoć analitičara podataka. Međutim, ljudsko oko može da posmatra ograničenu količinu dimenzija odjednom. Taj broj je uglavnom ograničen na tri dimenzije, prema tome ne mogu se otkriti neke kompleksne veze. Osim toga, analiza ovakvih veza bila bi vremenski zahtevna. Sa druge strane, rudarenje podataka je računarski proces koji otkriva obrasce u velikom skupu podataka koji uključuju elemente statistike, mašinskog učenja i veštačke inteligencije.

Izveštaji se generišu prilikom izvršavanja upita nad skladištem podataka. Upiti se postavljaju uz pomoć strukturiranog jezika upita, za strukturirane podatke. Podrazumeva se da skladište podataka ima definisanu šemu i da su podaci u uniformnom formatu. Uz pomoć njih moguće je prikazati uvide koji su otkriveni u podacima. Dakle, može se zaključiti da izveštaji pružaju informacije. Nasuprot tome, vizuelizacije se koriste za kreiranje značenja iz datih informacija. Skoro sva softverska rešenja poslovne inteligencije današnjice dolaze sa ugrađenom vizuelizacijom podataka. Vizuelizacije pružaju mogućnost automatske transformacije dobijenih rezultata u kružne grafike, stubičaste dijagrame, histograme, toplotne mape i druge vizuelne prikaze. Korisnicima je mnogo lakše na ovaj način da razumeju obrasce, odnose i trendove u poređenju sa rezultatima koji su prikazani u vidu tabele sa brojevima. Neka softverska rešenja uključuju i filtriranje, mešanje i druge manipulacije sa podacima radi dobijanja novih uvida i boljeg razumevanja podataka.

2.5. Sloj metapodataka

Mnogi autori opisuju metapodatke kao podatke o podacima. Drugim rečima, to su podaci koji korisnicima nekog objekta daju saznanja o postojanju objekta i o njegovim karakteristikama. U oblasti poslovne inteligencije oni opisuju izvore podataka, skladište podataka, poslovna pravila, autorizaciju pristupa i načine na koji se podaci izdvajaju, transformišu i obrađuju.

3. KORISTI IMPLEMENTACIJE BI SISTEMA

Implementacija softvera poslovne inteligencije i odgovarajućih skladišta/baza podataka je složen proces koji zahteva velika ulaganja novčanih i ljudskih resursa i vremena. Pre investicije neophodno je utvrditi da li je ona ekonomski opravdana.

Ulaganja ka potencijalnom rešenju poslovne inteligencije moraju biti opravdana potencijalnim koristima koja ta ulaganja mogu da donesu organizaciji.

Pored kvalitetnih vizuelnih prikaza koji omogućavaju korisnicima da lakše razumeju način na koji su podaci povezani, u nastavku će biti nabrojane konkretne koristi prilikom implementacije nekog rešenja poslovne inteligencije.

Odluke zasnovane na podacima: Kada su u pitanju taktičke i strateške odluke, menadžeri su prinuđeni da u velikoj meri budu oslonjeni na sopstvena znanja, iskustva i procene. Nepostojanjem BI softvera, oni se ne mogu osloniti na činjenice, te su u potpunosti prepušteni donošenju odluka koje su zasnovane na sopstvenom nagađanju. BI softver omogućava relevantne informacije koje pomažu da se donose odluke zasnovane na podacima. Kao rezultat, odluke koje se donose su kvalitetnije i smanjuje se neizvesnost u vođenju poslovanja.

Pristup ključnim informacijama: Velika količina podataka generiše se svakodnevno zahvaljujući velikom broju procesa u organizaciji i dinamičnom okruženju. Uz pomoć BI sistema moguće je skratiti vreme koje je potrebno za pronalaženje potrebnih informacije.

Takođe, zahvaljujući naprednim vizuelizacijama moguće je agregirati i lakše razumeti istorijske podatke ili podatke u realnom vremenu. Na osnovu istorijskih podataka moguće je praviti analize predviđanja koje utvrđuju dalje pravce poslovanja.

Povećana operativna efikasnost: BI alati pokrivaju celokupno poslovanje organizacije i mogu generisati izveštaje za svaki nivo u hijerarhijskoj strukturi organizacije. Zahvaljujući ovim izveštajima moguće je utvrditi nedostatke kao što su prazan hod, uska grla, nepotrebni troškovi i drugi nedostaci koji umanjuju vrednost ukupnog profita. Na osnovu ovih uvida menadžment može da optimizuje procese i ukloni nedostatke kako bi se povećala operativna efikasnost i profit organizacije.

Optimizacija resursa: BI alati pružaju informacije u formatu koji je potreban organizaciji, a kreiranje izveštaja je jednostavno, intuitivno i najčešće ne zahteva dodatne veštine za korišćenje alata.

Povećanje prodaje: Zahvaljujući podacima koji dolaze iz eksternih izvora, prvenstveno sa društvenih mreža i drugih interaktivnih platformi, organizacije mogu da prikupe ogromne količine podataka vezane za obrasce ponašanja kupaca prilikom kupovine. Takođe, analiziranje koje proizvode kupci gledaju na veb sajtu, koje poručuju i koje proizvode vraćaju mogu pomoći organizaciji da utvrdi dalje pravce delovanja

Segmentacija kupaca: Grupisanje kupaca na osnovu podataka o njihovim karakteristikama koje organizacije generišu zahvaljujući prijavljivanju prilikom kupovine ili popunjavanjem anketa. Ti podaci se mogu odnositi na starost, pol, adresu stanovanja, visinu prihoda, broj članova domaćinstva i slično. Dobijeni rezultati mogu se dalje koristiti kako bi se marketing kampanje prilagodile ciljnim kupcima.

4. ZAKLJUČAK

U dinamičnom i promenljivom okruženju današnjice, organizacije moraju da pronalaze načine kako da se brzo prilagode tržišnim trendovima i zahtevima kupaca. Pravovremene i efektivne odluke omogućavaju opstanak na tržištu i konkurentsku prednost. Alati poslovne inteligencije koriste se, ne samo za donošenje odluka podržanih podacima, već i za razumevanje tržišnih performansi organizacije i kreiranje strateških planova. Prema tome, poslovna inteligencija predstavlja ključnu komponentu u poslovanju organizacije ukoliko ona želi da zadrži svoju tržišnu poziciju, unapredi svakodnevno poslovanje i odredi ciljeve za budućnost koji će omogućiti rast i razvoj.

5. LITERATURA

- [1] Power, Daniel J. "A brief history of decision support systems." *DSSResources.com* 3 (2007).
- [2] Ranjan, Jayanthi. "Business intelligence: Concepts, components, techniques and benefits." *Journal of theoretical and applied information technology* 9.1 (2009): 60-70.
- [3] Niu, Li, Jie Lu, and Guangquan Zhang. "Business intelligence." *Cognition-Driven Decision Support for Business Intelligence*. Springer, Berlin, Heidelberg, 2009. 19-29.
- [4] Olszak, Celina M., and Ewa Ziemba. "Approach to building and implementing business intelligence systems." *Interdisciplinary Journal of Information, Knowledge, and Management* 2.1 (2007): 135-148.
- [5] Bălăceanu, Daniel. "Components of a Business Intelligence software solution." *Informatica Economică* 2.42 (2007): 67-73.
- [6] Ong, In Lih, Pei Hwa Siew, and Siew Fan Wong. "A five-layered business intelligence architecture." *Communications of the IBIMA* (2011).
- [7] William H. Inmon, "Building The Data Warehouse (4th Edition)", Wiley Publishing, Inc., 2005.

Kratka biografija:



Bojana Popov rođena je u Senti 1997. godine. Diplomski rad je odbranila na Fakultetu tehničkih nauka iz oblasti industrijsko inženjerstvo i inženjerski menadžment 2020. godine.