

UPRAVLJANJE ZNANJEM U SISTEMIMA POSLOVNE INTELIGENCIJE

KNOWLEDGE MANAGEMENT IN BUSINESS INTELLIGENCE SYSTEMS

Dr Ljubomir Trifunović, redovni profesor

ljubomir.trifunovic.efb@gmail.com

Ekonomski fakultet Brčko Univerziteta u Istočnom Sarajevu

Dr Lazar Radovanović, docent

lazar.radovanovic.efb@gmail.com

Ekonomski fakultet Brčko Univerziteta u Istočnom Sarajevu

Mr Slavko Matanović

"MySoftware", d.o.o. Brčko

slav.matan@gmail.com

APSTRAKT

U članku se obrađuju tehnologije sticanja i crpljenja podataka za potrebe izgradnje sistema poslovne inteligencije. Posebna pažnja posvećena je procesu upravljanja znanjem i otkrivanju znanja tehnikama rudarenja podataka iz skladišta podataka i na Internetu.

Ključne riječi: poslovna inteligencija, upravljanje znanjem, otkrivanje znanja, rudarenje podataka

ABSTRACT

The work deals with the technologies of acquiring knowledge and retrieving data from diverse data sources for the purpose of developing business intelligence systems. Knowledge discovery by means of data mining techniques used with various data formats and sources – data warehouse and the Internet – and the subsequent knowledge management are specially highlighted.

Key words: business intelligence, knowledge management, knowledge discovery, data mining

UVOD

U prvom dijelu članka istražuje se poslovna inteligencija, implementacija i područja primjene sistema poslovne inteligencije. Drugi dio rada obuhvata upravljanje znanjem u sistemima poslovne inteligencije, skladište podataka i alate za otkrivanje znanja iz skladišta podataka, posebno tehnike rudarenja (*data mining*) podataka.

POSLOVNA INTELIGENCIJA

Poslovna inteligencija predstavlja skup tehničkih i procesnih inovacija pomoću sistema *data warehousinga* i prostora poslovne inteligencije. Proaktivna poslovna inteligencija fokusira se na ubrzanju procesa donošenja odluka jačanjem infrastrukture poslovne inteligencije za identifikovanje, izračunavanje i trenutnu distribuciju informacija i znanja koji su relevantni za poslovanje preduzeća. Proaktivna poslovna inteligencija [Turban, E. et al., 2005, p. 107.] ima pet komponenti: (1) *real-time warehousing*, (2) automatsko otkrivanje anomalija i odstupanja, (3) proaktivno upozorenje pomoću automatskog primaoca rješenja, (4) neograničeno praćenje pomoću toka posla i (5) automatsko učenje i usavršavanje.

Pomoću sistema poslovne inteligencije preduzeće unapređuje poslovanje i upravljanje preko konkurenata istraživanjem i korišćenjem podataka o potrošačkim preferencijama, karakteristikama potrošača, lancima snabdijevanja, geografskim uticajima, formiranjem cijena i načina povećanja cjelokupne poslovne efikasnosti i efektivnosti [www.learnbi.com/bi.htm].

Implementacija sistema poslovne inteligencije u poslovni sistem

Prilikom usvajanja koncepta poslovne inteligencije preduzeće se može odlučiti za jednu od mogućih strategija uvođenja novog sistema. Pri tome treba imati na umu da, za razliku od sistema za obradu podataka, sistemi poslovne inteligencije rijetko su u potpunosti do kraja izgrađeni. Za ove sisteme podrazumijeva se

kontinuirani razvoj u skladu s poslovnom dinamikom preduzeća i promjenama u poslovnom okruženju. Glavni cilj razvoja sistema poslovne inteligencije ogleda se u postizanju optimalnog rješenja u određenom vremenskom periodu, uz uvažavanje činjenice da će se vremenom postojeći sistem nadograđivati kako bi kontinuirano zadovoljavao potrebe radi kojih se i uvodi [Howson, C., 2008, pp. 139-140].

Postoji nekoliko strategija uvođenja sistema poslovne inteligencije. Jedna od varijanti je da se promjene uvedu u one organizacione jedinice u kojima će se najprije vidjeti rezultati. Drugi pristup je da se promjene obave tamo gdje se očekuju najmanji otpori ili obrnuto, da se promjene prvo obave tamo gdje su najveći otpori, da bi se skeptici razuvjerali na samom početku uvođenja sistema. Bez obzira na izabranu strategiju, izgradnja sistema poslovne inteligencije predstavlja ogroman posao i to je, po pravilu, posao koji zahtijeva poseban, projektni pristup.

Prvi zadatak u razvoju projekta odnosi se na definisanje cilja koji se želi postići uvođenjem koncepta poslovne inteligencije. Jedan od glavnih zahtjeva je da se izgradi takav informacioni sistem pomoću kojeg je moguće brzo i efikasno stvoriti informacije i znanje potrebno za analizu nabavke, prodaje, troškova i novčanih tokova preduzeća. Zatim je potrebno definisati konkretna pitanja unutar navedenih funkcija koje treba riješiti. Na primjer, u slučaju prodaje, moraju se definisati proizvodi koji će se istraživati, potraživanja od kupaca po ročnosti, dani vezivanja i uplate kupaca, analiza bazne profitabilnosti itd. Moraju se definisati i osnovni zahtjevi korisnika, najčešće menadžera, koje će oni postaviti pred takav sistem, kao što su: brzina, jednostavnost rada i učenja, mogućnost korišćenja raznih izvora podataka i postojećih aplikacija, prezentacija podataka u poslovnim terminima, višedimenzionalni pristup, mogućnost prikazivanja više ili manje detalja, mogućnosti kompleksnih analiza i sl.

Prije implementacije potrebno je da se izvrši ocjena izvodljivosti predviđenog projekta, a zatim slijedi implementacija i postavljanje infrastrukture. Preporučuje se da se prvo započne s implementacijom jednostavne i skromne teme. Zatim se provodi edukacija krajnjih korisnika prezentacijom sadržaja i načinom služenja bazom podataka. Glavno obilježje sistema poslovne inteligencije je sposobnost prilagođavanja promjenama i novim uslovima u okruženju i preduzeću u koje se sistem uvodi. Sistemi poslovne inteligencije imaju sposobnost učenja prikupljanjem novih podataka iz eksternih izvora i njihovom kombinacijom sa iskustvom sistema stečenom iz ranijih obrada. Na taj način nastaju nove, kvalitetnije informacije i znanje. Djelotvornost sistema poslovne inteligencije raste s brojem iterativnih procesa, pri čemu sistem akumulira iskustvo kojim se koristi u budućim obradama. S obzirom na to da se svaki posebni slučaj prikupljanja i obrade podataka, uz sugerisanje rješenja, može odvijati posebnim iterativnim procesom, svaki sljedeći proces, zahvaljujući akumuliranom znanju i iskustvu, odvija se na kvalitetniji način.

Područja primjene sistema poslovne inteligencije

Organizacije koje u svojim poslovnim aktivnostima primjenjuju koncept poslovne inteligencije uspostavljaju bolju kontrolu nad informacijama, u odnosu na one koje ne koriste njene prednosti, te donose kvalitetnije odluke, stvaraju konkurentsku prednost i ostvaruju veći prihod. Mnoga srednja i velika preduzeća u svijetu već su ugradila i koriste se sistemom poslovne inteligencije za upravljanje važnim procesima, poput upravljanja nabavkom sirovina i materijala, osiguranja kvaliteta proizvoda i usluga, praćenja prodaje i sl. Zasad se taj sistem primjenjuje na nivou pojedinih organizacionih jedinica, odsjeka, odjeljenja itd. Međutim, naprednija i prodornija preduzeća ubrzano rade na standardizaciji sistema poslovne inteligencije na nivou cijelog preduzeća, da bi osigurala prednosti koje taj sistem donosi.

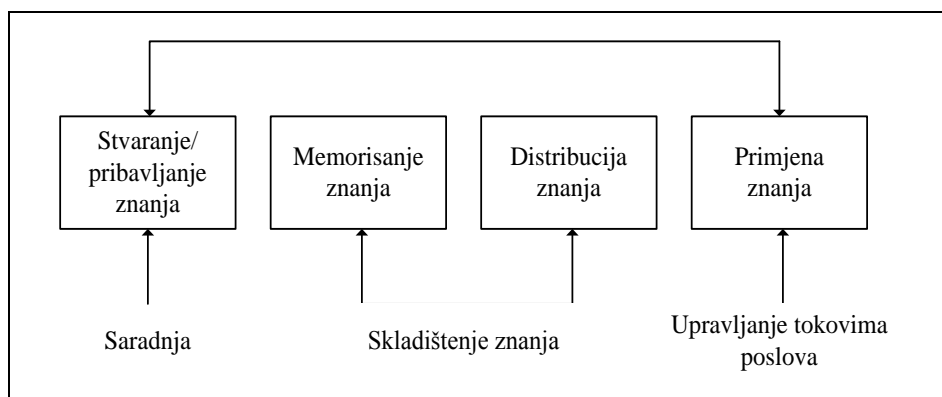
Porast konkurencije, zakonodavno uplitanje državnih vlasti i međunarodna standardizacija u mnogim područjima, zatim okrupnjavanje korporacija i sve veća prisutnost globalizacije u svim područjima, posljedica su pojačanog istraživanja u području poslovne inteligencije radi njene primjene i pomoći pri poslovnom odlučivanju. Koncept poslovne inteligencije danas se uspješno primjenjuje u mnogim poslovnim područjima i oblastima, a naročito u bankarstvu, maloprodajnim trgovinskim lancima, logističkim aktivnostima (transport, upravljanje zalihama), osiguravajućim društvima, farmaceutskoj industriji itd. Predviđa se, da će se prednostima koje pruža poslovna inteligencija u budućnosti koristiti i fizička lica – pojedinci da bi unaprijedili kvalitet svoga življenja. Praktična realizacija ove vizije je moguća zbog sve šire primjene današnje sofisticirane informacione tehnologije. Ovdje se, prije svega, misli na tehnologije *Web* usluga, ekstraneta i korporacijskih portala.

PROCES UPRAVLJANJA ZNANJEM U SISTEMIMA POSLOVNE INTELIGENCIJE

Naglašavanje uloge znanja posljedica je značajnih strukturnih promjena u razvoju ekonomije koja se, krajem prošlog vijeka, po svojoj suštini, potpuno mijenja tako da prelaskom iz industrijske u savremenu ekonomiju glavni faktor ekonomskih aktivnosti i produktivnosti postaje znanje. Prelaz iz industrijske ekonomije u ekonomiju znanja karakteriše nekoliko bitnih činjenica: informacija postaje važan resurs, a informaciono-komunikaciona tehnologija osnovna infrastruktura za sticanje i upravljanje znanjem. Razvojem te tehnologije znatno se ubrzava poslovanje i skraćuje proizvodni i poslovni ciklus, a odnos prema proizvodu mijenja se u korist potražne strane, odnosno kupaca.

Upravljanje znanjem predstavlja nezaobilaznu komponentu savremenog menadžmenta. Opšti cilj upravljanja znanjem je povezivanje onih subjekata kojima treba znanje s izvorima znanja u preduzeću i njegovom okruženju, kao i usklađivanje transfera znanja. Konkretni cilj upravljanja znanjem predstavlja proces utvrđivanja i analize raspoloživog znanja kako bi se ispunili postavljeni ciljevi poslovanja preduzeća. Upravljanje znanjem podrazumijeva konstantni proces obnove znanja čiji je cilj kontinuirana inovacija i modifikacija postojećeg znanja da bi preduzeće moglo da održava korak s dinamičkim tržišnim okruženjem.

Proces upravljanja znanjem sastoji se od nekoliko faza (slika 1): stvaranje, odnosno pribavljanje znanja, memorisanje, distribucija i primjena znanja. U organizaciji koja uči (*learning organization*) cilj je kontinuirana obnova znanja pa je ovaj proces stalan i cikličan.



Slika 1. Ključne faze procesa upravljanja znanjem

Izvor: [www.brint.com/casestudies]

U današnje vrijeme preduzeća koja žele da postanu i ostanu tržišni lideri treba da vode računa o sopstvenoj infrastrukturi upravljanja znanjem. Pri tome potreba za novim znanjem mora da bude jedan od strateških prioriteta poslovanja, koji prepoznaje i prihvata najviši nivo menadžmenta. U tom smislu potrebno je izgraditi organizacionu kulturu i razvijati mehanizme za distribuciju znanja. Nakon stvaranja okvira koji podstiče razmjenu znanja potrebno je analizirati i definisati jaz između raspoloživog i potrebnog znanja, a nakon toga razvijati mehanizme procjene i reprodukcije znanja. U tom smislu preduzeća moraju pronaći metode i tehnike otkrivanja, prikupljanja, memorisanja, reprodukcije i dopunjavanja znanja. Krajnji cilj je postizanje sinergijskih efekata kombinacijom procesiranja podataka i informacija uz pomoć informacionih tehnologija s kreativnim i inovativnim sposobnostima zaposlenih. Ovdje se govori o lancu vrijednosti koji podrazumijeva stalnu transformaciju sirovih podataka u informacije i znanje, pri čemu je krajnji cilj donošenje kvalitetnijih poslovnih odluka [www.brint.com/casestudies].

Značaj upravljanja znanjem

Konkurentska prednost organizacije zavisi od uspješnosti upotrebljavanja znanja a naročito od brzine usvajanja inovacija. U žestokoj konkurentskoj borbi važno je pronaći i zadržati neiskorišćeno znanje u organizaciji i pretvoriti ga u intelektualni kapital. Sistematsko prikupljanje i upotreba znanja radi generisanja novih ideja postiže se stvaranjem odgovarajućeg okruženja u kojem upotreba novih informaciono-komunikacionih tehnologija znatno olakšava savladavanje toga problema.

Sposobnost kontinuiranog generisanja novih ideja i stvaranje inovativnih proizvoda karakteristika je uspješnih organizacija za koje inovacija predstavlja dio strategije poslovanja i koje posjeduju odgovarajuću plitku i savremenu organizacionu strukturu i menadžment koji podstiče inovacije. Osnov kreativnosti i inovativnosti postiže se i stvaranjem timova od izvršilaca različitih profila, što je preduslov za kreiranje i integrisanje različitog znanja i generisanje novih ideja. S obzirom na to da se do inovacije, po pravilu, ne dolazi za kratko vrijeme bitno je da menadžment ne očekuje brze rezultate, već da sistematski usmjerava i razvija definisanu politiku te podržava aktivnosti bitne za praktično provođenje upravljanja znanjem.

SAVREMENE INFORMACIONE TEHNOLOGIJE I INTERNET U ELEKTRONSKOJ DIMENZIJI UPRAVLJANJA ZNANJEM

Razvoj savremene informacione tehnologije i Interneta omogućili su reorganizovanje kvalitativnih parametara: ideja, inovacija, intelektualnog kapitala i znanja, te povezivanje s kupcima zadovoljavanjem njihovih potreba proizvodima i uslugama na kvalitetniji način. Glavne prednosti Interneta kao globalne komunikacione mreže ogledaju se u savladavanju prostornih ograničenja, ubrzavanju komunikacionih procesa i brzom i efikasnom snabdijevanju relevantnim stručnim i opštim znanjem koje je potrebno za donošenje poslovnih odluka.

Savremene informacione tehnologije i Internet povećali su potrebe i snagu upravljanja znanjem. Prednosti koje oni daju ogledaju se u jednostavnijoj distribuciji znanja unutar i izvan poslovnih sistema. Danas ta tehnologija stvara sredstva koja omogućavaju prikupljanje i skladištenje znanja i stvaranje novih znanja neophodnih za poslovno odlučivanje. Informaciono-komunikaciona tehnologija još ne može u potpunosti zamijeniti vrijednost i potrebu direktne komunikacije, naročito kada se radi o distribuciji iskustvenog znanja, ali može uspješno asistirati u posredovanju i olakšavanju kreiranja mreže koja se zasniva na znanju. Ova tehnologija, posebno infrastruktura Interneta (naročito *Web* i *e-mail*), ima naročitu vrijednost i igra ulogu posrednika za ona preduzeća koja posluju na udaljenim i različitim geografskim lokacijama i gdje je onemogućena direktna komunikacija između tražilaca znanja i onih koji ga nude.

Razvoj Interneta podloga je za razvoj globalnog tržišta znanja na kojem se susreću ponuđači i kupci znanja. Informacije danas postaju sve dostupnije, pa se povećava i uticaj pojedinaca na poslovanje organizacija. Znanje pojedinca ima snažan uticaj na preduzetništvo koje se oslanja na znanje. Taj proces nadopunjen je i Internetom. Donedavno su velika preduzeća održavala svoju moć upravljanjem informacijama, u okviru i izvan svojih granica, a danas to čine razmjenom i distribucijom znanja. Za razliku od drugih roba koje se razmjenjuju na berzama, prednost i karakteristika znanja jeste da je ono obnovljiv izvor, te da se ne troši, već, naprotiv, upotrebom znanje se akumulira, a njegova vrijednost raste.

Internet je sredstvo koje pruža uslugu jednostavne i sveobuhvatne razmjene znanja i oslobađanje vrijednosti i potencijala ljudskih mozgova. Znanje i učenje danas postoje sve više dio organizacione kulture i organizacionih procesa. Procesi, proizvodi i usluge postaju intenzivni znanjem. Preduzeća koja dobro ne upravljaju svojim znanjem marginalizovana su i prijeti im neuspjeh na tržištu i propadanje.

Organizacije koje uče u prvi plan stavljaju znanje i zajedničko učenje kao osnovu svoje budućnosti. Međutim, javlja se problem stvaranja preduslova i infrastrukture potrebne za upravljanje poslovnim znanjem i stvaranje organizacije koja uči. Razvoj informaciono-komunikacione tehnologije i inteligentnih i drugih sistema upravljanja znanjem omogućava uspješno rješavanje navedenih problema. Veliki i značajan podskup upravljanja poslovnim znanjem i prvi korak prema organizaciji koja uči jeste grupa metoda, alata i aplikacija poslovne inteligencije.

SKLADIŠTA PODATAKA

Skladište podataka (*data warehouse*) podrazumijeva skup podataka izolovanih iz operativnih baza i spremljenih u posebne baze. Glavna karakteristika koja determiniše skladište podataka je njegova namjena. U skladištu podataka podaci se prikupljaju i organizuju na način da budu lako dostupni menadžmentu kojima se on može koristiti na brz i jednostavan način za potrebe analize poslovanja preduzeća i donošenje poslovnih odluka. Skladište podataka predstavlja tematski, odnosno predmetno orijentisan, integrisan i sadržajno nepromjenljiv i vremenski zavisn skup podataka čiji je cilj podrška donošenju odluka.

Predmetna orijentisanost podataka znači da se oni organizuju predmetno, odnosno tematski na način da daju informacije o tačno određenim temama u okviru funkcionalnih područja, na primjer u okviru marketinga, prodaje, nabavke itd., umjesto o tekućim poslovnim aktivnostima preduzeća. Suprotno tome operativne baze podataka organizuju se oko poslovnih aplikacija, dakle usmjerene su na tekuće operacije (npr. obrade narudžbi, isporuka i sl.). Integrisanost znači da se podaci prikupljaju u bazu podataka iz različitih izvora i skladište uvijek u istom formatu, te su konzistentni i prikazuju se na dosljedan način. Vezanost uz vrijeme – svi podaci u skladištu podataka vezani su i identifikuju se uz određeni vremenski period, što znači da imaju istorijski karakter. Za razliku od njih, u operativnim bazama podataka smješteni su samo aktuelni, najsvježiji podaci. Međutim, sa gledišta koncepta poslovne inteligencije, sveobuhvatno predviđanje budućih događaja nije moguće provesti bez poznavanja istorije istih ili drugih događaja. Iz toga slijedi, bez obzira na to što podaci u skladištu podataka odražavaju prošlost, da je njihova usmjerenost na budućnost. Sadržajna nepromjenljivost – podaci u skladištu su stabilni i kad se jednom memorišu u skladište u pravilu se ne mijenjaju. Time se omogućava da donosilac poslovne odluke bude siguran da će dobiti jednak odgovor nezavisno od vremena ili učestalosti postavljanja upita.

Postupak skladištenja podataka predstavlja kontinuirani proces planiranja, kreiranja i prikupljanja podataka iz različitih izvora te korišćenja, održavanja, upravljanja i stalnog unapređivanja skladišta podataka. Između mnogih koraka u tom kompleksnom kontinuiranom procesu bitno je naglasiti važnost posjedovanja vizije o tome šta se želi postići kreiranjem skladišta podataka. Jedna od uloga skladišta podataka je razvijanje i korišćenje znanja zasnovanog na podacima i informacijama [<http://www.intranetjournal.com>]

Aktivnost skladištenja podataka predstavlja kontinuirani proces, a investicija uvođenja skladišta podataka je skupa i dugotrajna. Prilikom procesa donošenja odluke o kreiranju i implementaciji skladišta podataka potrebno usaglasiti i niz pitanja bitnih za uspostavljanje projekta skladišta. Prije implementacije projekta potrebno je identifikovati poslovni interes za izgradnjom i upotrebom skladišta podataka za potrebe poslovanja, dogovoriti izvore finansiranja, razviti kriterijume za određivanje poslovne upotrebljivosti skladišta podataka, provesti intervjuisanje korisnika o traženim informacijama, identifikovati izvore podataka za popunjavanje skladišta podataka, odlučiti se za veličinu skladišta podataka, utvrditi vrstu s aspekta sadržaja, odrediti fizičku lokaciju, donijeti odluku o izgradnji ili kupovini skladišta podataka, izabrati najpovoljnije alate i sisteme za upravljanje bazama, riješiti pitanje zapošljavanja itd. Na kraju implementacije slijedi puštanje sistema u rad, trening korisnika radi potpunog iskorišćavanja instaliranih alata, upravljanje sistemom skladištenja dodavanjem, modifikovanjem i njegovim razvijanjem.

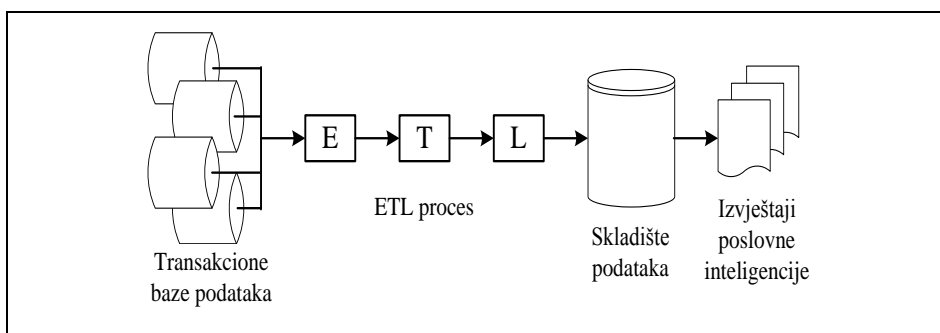
Glavni cilj skladišta podataka je oslobađanje informacija koje su zarobljene u operacionim bazama podataka i njihovo kombinovanje s informacijama iz ostalih, po pravilu, eksternih izvora podataka. Mnogi poslovni i drugi sistemi sve više traže dodatne podatke iz eksternih izvora, kao što su npr. podaci o konkurenciji, demografski podaci, prodajni trendovi i sl. jer kvalitetno donošenje poslovnih odluka zahtijeva pristup raznim izvorima podataka. Uspostavljanjem skladišta podataka operativne baze podataka rasterećuju se složeni upiti, pa dolazi do unapređenja njihovih operativnih funkcija. Na taj način iz operativnih baza uklanja se ogromna količina istorijskih podataka i preseljava u skladište podataka. Informacioni sistem, u tom slučaju, sastoji se od dva dijela, operativnog dijela i skladišta podataka. Budući da se tako lakše kontroliše i restrukturira informacioni sistem postaje efikasniji i efektivniji.

Pomoću tehnika otkrivanja znanja skladište podataka osigurava stalno pronalaženje novih informacija, zavisno od novonastalih uslova i zahtjeva. Procesi za generisanje informacija (ekstrakcija, agregacija, analiza, izvještavanje i itd.) izdvojeni su iz operativnih procesa, što znači da operativni nivo sistema više nije njima opterećen. Skladište podataka postaje mjesto prikupljanja i memorisanja poslovnih podataka i izvor informacija i znanja potrebnih pri poslovnom odlučivanju. Strukturu skladišta podataka čine dva osnovna dijela: podaci i mehanizmi manipulacije tim podacima. Dio skladišta u kojem se nalaze podaci sastoji se od osnovnih podataka i agregiranih višedimenzionalnih podataka, a mehanizme manipulacije predstavljaju procesi ekstrakcije, transformacije i punjenja podataka, sistem upravljanja podacima, postupci analitičke obrade i prezentacija rezultata.

Osnovna funkcija skladišta podataka je prikupljanje podataka i stvaranje logički integrisanih i predmetno usmjerenih informacija. Treba ga oblikovati tako da se može na jednostavan i brz način prilagođavati svim

promjenama i zahtjevima poslovnog okruženja. S obzirom na predmetnu usmjerenost podataka, pri modeliranju skladišta primjenjuju se tehnike koje podržavaju takvu orijentaciju i obezbjeđuju dovoljnu prilagodljivost da bi se vremenom mogli integrisati i podaci iz mogućih dodatnih izvora. Skladište podataka mora da bude izvor stabilnih podataka, nezavisnih od eventualnih promjena u poslovnim procesima. Iz tog razloga potrebno je koristiti model neosjetljiv na uticaje operativnih procesa koji kreiraju većinu podataka. Oslobođeno operativnih obrada, skladište podataka obezbjeđuje unapređenje procesa generisanja informacija tehnikama otkrivanja znanja i stalno pronalaženje novih informacija i kreiranja znanja.

U skladište podataka smještaju se podaci iz različitih izvora, najčešće iz transakcionih sistema organizacija. Najopsežniji posao u aktivnostima skladištenja podataka predstavljaju procesi integrisanja podataka i organizovanje njihovog sadržaja. Pri tom glavnu ulogu predstavlja skup procesa čiji je zadatak izdvajanje, transformisanje i punjenje ili unošenje podataka iz jednog ili više transakcionih sistema u skladište podataka (slika 2). Zajednički naziv za te procese je ETL proces, nastao od prvih slova engleskih riječi *extraction*, *transformation* i *loading* [Radovanović, K. L., 2012, str. 45-46].



Slika 2. ETL proces

Izvor: [Radovanović, K. L., 2012, str. 46.]

Prije ETL procesa potrebno je izvršiti pripremne aktivnosti vezane uz reformatiranje, usklađivanje i čišćenje podataka. Izvorne podatke, preuzete iz različitih datoteka i baza podataka, potrebno je unificirati, odnosno prikazati u jedinstvenom formatu u kojem će se podaci koristiti u svim daljnjim fazama obrade. Usklađivanje podataka provodi se da bi se izbjegla njihova redundancija.

Pored toga što se u informacionom sistemu isti podaci mogu pojaviti na više mjesta, oni mogu da budu nekonzistentni, odnosno da njihove vrijednosti nisu iste na svim mjestima na kojima se pojavljuju. Radi toga potrebno je takve podatke otkriti i uskladiti. Čišćenje, pripremna aktivnost ETL procesa, ima zadatak da ukloni podatke koji se pojavljuju kao posljedica ranijih grešaka u radu informacionih sistema ili zbog namjernog unošenja netačnih i lažnih podataka u sistem. Opšta karakteristika standardnih ETL alata je da ako imaju bolje performanse vezane uz procese čišćenja tada im je slabiji kapacitet procesa transformacije i obrnuto. Stoga je prije izbora potrebno poznavati karakteristike podataka koji će se smještati u skladište podataka. Ako se unaprijed zna da će biti mnogo podataka koji zahtijevaju postupak transformacije, tada se moraju odabrati ETL alati koji su efikasniji za proces transformacije i obrnuto.

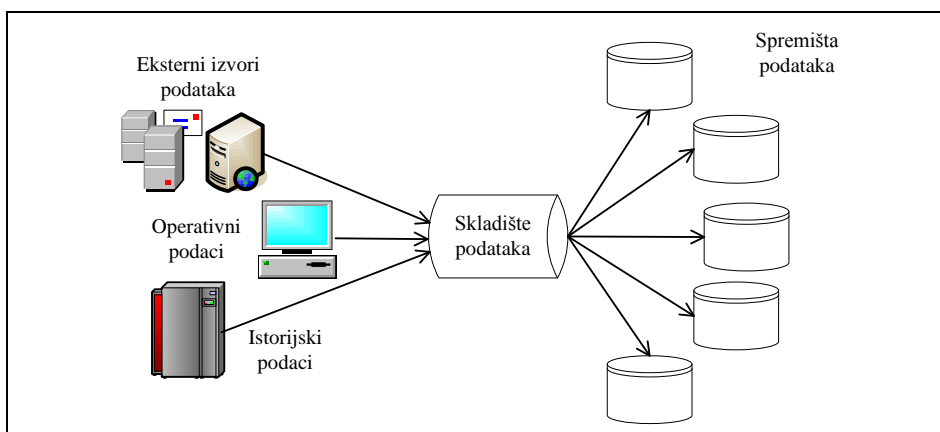
Proces ekstrakcije podataka potrebno je provoditi na način da pri tome redovni operativni poslovi što manje trpe. Stoga su programi i alati za ekstrakciju oblikovani tako da ETL procese mogu obavljati što efektivnije i brže. Pri tome se kao problem može pojaviti potencijalno visok stepen redundancije podataka u transakcionim sistemima, pa treba odabrati takav pristup ekstrakciji kojim se zahvataju samo oni podaci kojim će se koristiti aplikacije poslovne inteligencije. Podaci u transakcionim sistemima nalaze se u bazama podataka koje obuhvataju kraće vremensko razdoblje pa, zato, nisu dovoljni kao podloga za kvalitetne analize i podršku odlučivanju. Pored njih, dio izvornih podataka koji se obrađuje ETL procesima čine i istorijski podaci. Zato se za procese punjenja skladišta podataka koristi više vrsta ETL programa kao što su programi za inicijalno punjenje, programi za punjenje istorijskih podataka i programi za postupno (inkrementalno) punjenje.

Karakteristika programa za inicijalno punjenje skladišta podataka jeste da sadrže rutine za čišćenje i usklađivanje podataka, da bi se iz podataka eliminisale greške. Za istorijske podatke ponekad nije moguće primijeniti postupke čišćenja koji se primjenjuju za tekuće podatke, jer je od vremena nastanka tih podataka

do danas možda došlo do različitih promjena u slogovima i formatima podataka. Zato se za tu kategoriju podataka primjenjuju programi za punjenje istorijskih podataka koji čine nastavak inicijalnog punjenja. Za razliku od tekućih, istorijski podaci su statičkog karaktera i čine samo sadržaj arhivskih datoteka. Treću vrstu predstavljaju programi za inkrementalno punjenje podataka, a aktiviraju se nakon što su prethodna dva programa obavili postupak čišćenja i usklađivanja podataka. Njihova karakteristika je da se pokreću periodično, a predstavljaju stalno aktivan mehanizam punjenja skladišta podataka odgovarajućim sadržajima [Radovanović, K. L., 2012, str. 47].

Implementacija skladišta podataka u poslovni sistem

Razvoj i implementacija sistema skladišta podataka zahtijeva dosta vremena i značajna finansijska sredstva. Međutim, kako su danas uspješne organizacije svjesne činjenice da posjedovanje pravih i pravovremenih informacija predstavlja stratešku imovinu, odnosno mogućnost brze reakcije na stanje na tržištu, odlučuju se na korišćenje tehnologije skladišta podataka jer je upravo ona preduslov za brzo i fleksibilno pretraživanje.



Slika 3. Troslajni model skladišta podataka

Izvor: [Radovanović, K. L., 2012, str. 47.]

Jedan od praktičnih pristupa implemetacije sistema je postepena izgradnja uz sukcesivno nadograđivanje. Ukoliko se primijeni taj pristup, završni korak predstavlja kreiranje trosljednog modela sistema skladištenja podataka (slika 3). Postupak započinje razvojem nekoliko spremišta podataka koja podržavaju troslajnu arhitekturu. Nakon određenog vremena, kad je već razvijeno nekoliko spremišta podataka, slijedi izgradnja centralnog skladišta i odvajanje od pojedinih spremišta podataka. Postepena izgradnja i implementacija sistema skladišta podataka predstavlja prihvatljivo rješenje za većinu organizacija s obzirom na to da je moguće vremensko razgraničenje troškova.

Uvođenjem i puštanjem u rad ovog sistema završava se ogroman posao. Međutim, to nije kraj. Naime, implementacijom skladišta podataka i njegovim korišćenjem u praksi, mogu se pojaviti nova pitanja i zadaci s pozitivnim i negativnim predznakom koje treba riješiti. Na primjer, u postupku skladištenja podataka i korišćenja dobijenih informacija, može se pojaviti potreba za podacima koje nije moguće dobiti iz postojećih transakcionih sistema. U takvim slučajevima, za dopunu tim podacima, pojaviće se potreba za nadogradnjom ili modifikacijom sistema za transakcionu obradu podataka ili razvoja sistema namijenjenog za dobijanje podataka koji nedostaju. Dok tradicionalni projekti počinju postavljanjem zahtjeva, a završavaju podacima, projekti skladišta podataka počinju definisanjem podataka, a završavaju zahtjevima.

Višedimenzionalni prikaz podataka

Skladište podataka puni se izvorima podataka iz transakcione baze podataka preduzeća i eksternim izvorima podataka. Postupcima ekstrakcije, transformacije i punjenja, iz unutrašnjih i spoljnih izvora, dobijaju se osnovni podaci skladišta. Primjenom sistema za upravljanje podacima nastaju agregirani, višedimenzionalni podaci koji analitičkim metodama obrade omogućavaju dobijanje različitih informacija potrebnih menadžerima u procesu donošenja odluka. Najveća prednost dimenzionalne strukture je mogućnost vizualne prezentacije informacija. Povezivanje skladišta podataka s korisničkim interfejsom, uz određivanje načina postavljanja upita i prikaza rezultata upita, omogućeno je menadžerima jednostavno i brzo postavljanje upita

kao i odgovarajući prikaz rezultata upita i to prema potrebi u obliku teksta, slike, grafikona i sl. Zato se u skladištu podataka najčešće koristi dimenzionalna struktura podataka koja se zasniva na elementima relacije baze podataka, odnosno na relacionim tabelama [Radovanović, K. L., 2012, str. 49].

Vežni element koji objedinjuje sve elemente predstavlja središnja relaciona tabela koja memoriše elemente primarnih šifri svake od relacionih tabela koje učestvuju u strukturi, te numeričke varijable koje treba da se analiziraju [www.dwininfocenter.org]. Višedimenzionalna struktura podataka pruža velike mogućnosti za otkrivanje mnogih detalja različitim postupcima analitičke obrade kao što su agregacija i detaljizacija ili svrdlanje (*drill up, drill down, drill through*), unakrsni tabelirani prikaz (*cross tabulation*), selekcija, isijecanje, izdvajanje i kombinovanje svih dimenzija, rotacija, odnosno isticanje jedne dimenzije dok su druge u pozadini (*pivoting*), prognoziranje, modeliranje, grafičko prikazivanje (*charting*), statističke analize (trend, klasteri) itd.

Navedene tehnike otkrivanja znanja omogućavaju kontinuirano pronalaženje novih informacija namijenjenih menadžerima, a služe im za strateško, taktičko i operativno donošenje odluka. Uvođenjem koncepta skladištenja podataka, operativne baze podataka prestaju da budu opterećene složenim upitima, pa cijeli informacioni sistem, koji se sada sastoji od dva dijela – operativnog i skladišta podataka, postaje efikasniji i lakše se kontroliše i restrukturira.

Otkrivanje znanja data mining tehnikama

Kontinuirani proces informatizacije društva, zapisivanje podataka, teksta i drugih sadržaja u digitalnom obliku, te stalno memorisanje podataka u baze podataka dovodi do njihovog izuzetno velikog rasta. Istovremeno s rastom baza podataka razvija se i potreba da se velike količine podataka memorisane u njima, analiziraju i vizualiziraju, da bi se došlo do potrebnih podataka, informacija i znanja. Pravi podaci i informacije predstavljaju osnovu za donošenje uspješnih poslovnih odluka. U nastavku teksta opisan je koncept rudarenja podataka koji se primjenjuje na području skladištenja podataka, pretraživanja podataka i otkrivanja znanja u bazama podataka, s naglaskom na poslovni, a manje na tehnički aspekt.

Obično se na rudarenje podataka (*data mining*) gleda kao na finalnu manifestaciju procesa skladištenja podataka. Međutim, danas već dolazi do izdvajanja u potpodručja koja se povezuju na izvore podataka pa se govori o rudarenju teksta (*text mining*), Web-a (*Web mining*) ili podataka organizovanih u vremenske serije.

Rudarenjem podataka pronalaze se skriveni trendovi, modeli, zakonitosti i odnosi među podacima. Alati za rudarenje podataka daju odgovore na poslovna pitanja za čija je rješavanja u tradicionalnom pristupu trebalo mnogo više vremena. Oni pretražuju baze podataka tražeći skrivene modele i pronalaze predvidljive informacije koje mogu da promaknu stručnjacima jer "leže" izvan njihovih očekivanja. Primjenom alata za rudarenje podataka moguće je otkriti ranije neprepoznatljive matrice ponašanja, mogu se tačnije i brže predvidjeti budući trendovi i ponašanja, što omogućava poslovnom svijetu donošenje proaktivnih odluka zasnovanih na znanju (*knowledge-driven decisions*). Razvitak metoda koje se danas koriste u okviru pojma rudarenja podataka započinje sedamdesetih i osamdesetih godina prošlog vijeka, a od sredine devedesetih godina pojam rudarenje podataka objedinjuje skup metoda i postupaka čiji je zajednički cilj otkrivanje zakonitosti u masi podataka.

Tehnike rudarenja podataka rezultat su procesa istraživanja i razvoja statističkih algoritama. Ova evolucija započela je još kad su poslovni podaci prvi put memorisani u računare, a nastavlja se kontinuirano unapređenjem pristupa podacima, generisanjem tehnologija koje omogućavaju korisnicima navigaciju kroz podatke u realnom vremenu. Proces rudarenja podataka danas je moguće provoditi zato što je potpomognut: moćnom multiprocesorskom računarskom tehnologijom, tehnologijom za masivno prikupljanje podataka i algoritamskim tehnikama za rudarenje podataka. U evoluciji od poslovnih podataka do poslovnih informacija i znanja, svaki novi korak gradi se na prethodnom. Na primjer, dinamični pristup podacima je kritična tačka za svrdlanje (*drill-through*) u aplikacijama za navigaciju podacima, a osposobljenost za skladištenje ogromnih baza podataka je kritična za proces rudarenja podataka.

U mnogim poslovnim područjima, na primjer finansijama, trgovini, marketingu, zdravstvu ili u bilo kojem naučnoistraživačkom području, standardni pristup analizi podataka zasniva se na radu analitičara koji

obrađuju podatke uz primjenu računarskih programa ili bez njih. S obzirom da količina podataka znatno raste, danas je bilo kakva obrada podataka bez upotrebe računara i savremenih algoritama obrade potpuno neefikasna i praktično nemoguća. U okviru razvoja informacionih sistema, kao što je ranije napomenuto, prisutna je pojava rasta baza podataka, koja je posljedica kontinuiranog memorisanja podataka u baze podataka. Savremena analiza takvih podataka preduslov je za donošenje kvalitetnih odluka u modernom poslovanju, a isto tako i kvalitetnog naučnoistraživačkog rada. Rezultat analize podataka iz baza je otkrivanje novog znanja. Izraz otkrivanje znanja podrazumijeva cjelokupni proces otkrivanja korisnog znanja iz podataka, a rudarenje podataka je samo jedan korak u tom procesu. U procesu rudarenja podataka koriste se različite metode, međutim analitička statistika predstavlja temelj svih postupaka otkrivanja znanja. Gledano iz perspektive statistike, postupak rudarenja podataka uz pomoć računara je automatizovana istraživačka analiza podataka iz velikih i kompleksnih baza podataka.

Uspješnost rudarenja podataka zavisi od izvora podataka i od samog kvaliteta podataka. U organizacijama koje su razvile i koriste se alatima poslovne inteligencije izvori podataka za analizu najčešće su skladišta podataka. Međutim, postojanje skladišta podataka ne mora da bude preduslov za proces rudarenja podataka. Kreiranje i izgradnja skladišta podataka često je ogroman posao koji ponekad traje godinama i pretpostavlja znatna finansijska sredstva. Zato kao izvori podataka ponekad služe i druge, izvorne baze podataka u kojima podaci mogu biti nekompletni, ne pojavljuju se na zadovoljavajućem stepenu granulacije ili se može pojaviti nekonzistentnost unutar samih podataka.

Metode data mininga

Metode rudarenja podataka mogu se primijeniti u područjima u kojima se raspoloživo velikom količinom podataka na osnovu kojih se otkrivaju određene veze, pravilnosti i zakonitosti. Za uspješno provođenje procesa rudarenja podataka potrebno je precizno formulisati cilj i problem koji se želi riješiti, a uspješnost rudarenja podataka zavisiće i od kvaliteta raspoloživih podataka. Kada se radi o izboru metoda rudarenja podataka, danas na raspolaganju stoji veliki broj različitih metoda. Postoje glavne, opšteprihvaćene metode rudarenja podataka, dok, isto tako, postoji i niz metoda iz ostalih područja koje se ne mogu svrstati u jednu vrstu metoda. Metode koje se najčešće koriste u procesu rudarenja podataka su: regresiona metoda, klasifikacione metode, metode klasterisanja, metode koje se zasnivaju na vještačkim neuronskim mrežama, stabla odlučivanja, genetski algoritmi i dr.

Nakon provedenog postupka rudarenja podataka u kojem dolazi do otkrivanja niza korisnih pravila, da bi se otkrivena pravila mogla uspješno koristiti i interpretirati, potrebno ih je povezati i formalizovati. Znanja otkrivena u procesu rudarenja podataka prezentiraju se u obliku izvještaja ili se formaliziraju u "skladište" u sisteme zasnovane na pravilima. Skladištenje se primjenjuje za ona pravila koja se izražavaju u obliku *ako – onda* (*if – then*). Neki sistemi, kao npr. neuronske mreže, u mogućnosti su, osim pravila, da prihvate dinamičke modele. Sistemi zasnovani na pravilima (*rule based systems*) mogu se podijeliti na dva osnovna tipa: tradicionalni ekspertni sistemi i sistemi zasnovani na rasplinutoj (*fuzzy*) logici. Gledano iz korisničke perspektive, ekspertni sistemi djeluju po principu pitanja i odgovora. Mehanizam odgovora je složen proces koji ulančava niz pravila radi zaključivanja na osnovu odgovora koje mu nudi korisnik. Strukturirano znanje u obliku pravila može se dalje nadopunjavati, odnosno referisati na "tvrda pravila" već postojeće baze činjenica ekspertnog sistema. Na taj način dolazi do interakcije dijelova ekspertnih sistema i novih pravila dobijenih različitim metodama rudarenja podataka [Radovanović, K. L., 2012, str. 52-53].

Primjena rudarenje podataka

Danas postoji velik broj organizacija koje nude softver za rudarenje podataka, s tim da je manji broj onih koje nude kompletna rješenja za otkrivanje znanja. U današnjim uslovima poslovni procesi generišu ogromne količine podataka, a jaz između mogućnosti prikupljanja i analize podataka stalno se povećava. Unutar te mase podataka skrivaju se informacije i znanje koje je od strateškog značaja za poslovne subjekte. Dva kritična faktora bitna za uspješno provođenje rudarenja podataka su pravilno integrisano skladište podataka i dobro poznavanje i razumijevanje procesa poslovanja nad kojim se želi primijeniti postupak rudarenja podataka. Naravno, primjena rudarenja podataka podrazumijeva dobro statističko predznanje, a većina preduzeća, obično, nema takvog analitičara ili nema dovoljan broj analitičara koji se time bave.

Data mining najkorisniji je tamo gdje je stalno prisutna prijetnja viška podataka. Rudarenje podataka primjenjuje se na mnogim poslovnim područjima kao što su proizvodnja, telekomunikacije, bankarstvo, finansije, osiguranje, maloprodaja i dr. Situacije u kojima se rudarenje podataka uspješno primjenjuje su npr. zadržavanje potrošača, otkrivanje prevara s kreditnim karticama, sistemi za segmentaciju, targetiranje i pozicioniranje na tržištu i sl. Na primjer, finansijska institucija koja je izložena rastućoj konkurenciji zbog internacionalne deregulacije posla, te zbog toga gubi sve veći broj korisnika, rudarenjem podataka po internim transakcionim bazama podataka i ostalim spoljnim izvorima, može da sazna više o razlozima zbog kojih njeni korisnici odlaze konkurenciji. Na osnovu novih informacija i znanja dobijenog postupkom rudarenja podataka može se lakše doći do saznanja o načinima za ublažavanje navedenog problema.

ZAKLJUČAK

Razvojem i primjenom sistema poslovne inteligencije pruža se mogućnost organizaciji da se koristi istorijskim podacima iz baza i skladišta podataka za generisanje informacija i znanja iz skrivenih struktura podataka na osnovu kojih se mogu analizirati prošla kretanja i iz toga izvući zaključci i predvidjeti budući pravci razvoja. Tehnikama rudarenja podataka (*data mining*) i novim pristupom dinamičnom istraživanju podataka iz velike količine podataka dolazi se do skrivenih struktura podataka koje su značajne za pribavljanje informacija i otkrivanje znanja kojim se upravlja da bi se stvorila nove poslovne vrijednosti.

Principi nove ekonomije zasnivaju se na znanju i učenju o načinu sticanja i upravljanja znanjem. Oštra konkurencija na globalnom tržištu prisiljava organizacije na odbacivanje uhodanih stereotipa i tradicionalnog načina poslovanja i na stvaranje novog pozitivnog nadahnuća i kreativnog radnog okruženja koje podstiče zaposlene na razmjenu ideja i znanja i stvaranje organizacione kulture koja se zasniva na učenju. Savremene organizacije pretvaraju se u tvornice znanja, njihovi zaposleni u kreatore znanja, a upravljanje znanjem postaje osnovni organizacioni model savremenog poslovanja. Savremena informaciona tehnologija, sistemi poslovne inteligencije i Internet omogućavaju bolje upravljanje i bržu distribuciju znanja.

LITERATURA

- [1] Turban, E. et al. (2005). *Decision Support Systems and Intelligent Systems*. N.J.: Pearson Prentice Hall.
- [2] www.learnbi.com/bi.htm.
- [3] Howson, C. (2008). *Successful Business Intelligence: Secrets to Making BI a Killer App*. McGraw-Hill.
- [4] www.brint.com/casestudies, "Knowledge in Action: Case Studies (Worldwide)".
- [5] <http://www.intranetjournal.com/features/datawarehousing.html>.
- [6] Radovanović, K. L. (2012). *Sistemi podrške odlučivanju*, Ekonomski fakultet, Brčko.
- [7] www.dwinifocenter.org/gotch.html, *Data Warehousing Gotchas*.