



## Osnovni koncepti Data Warehouse sistema

---

*Sistemi skladišta podataka*

# Sadržaj

- Motivacija nastanka DW sistema
- Konceptija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Motivacija nastanka DW sistema

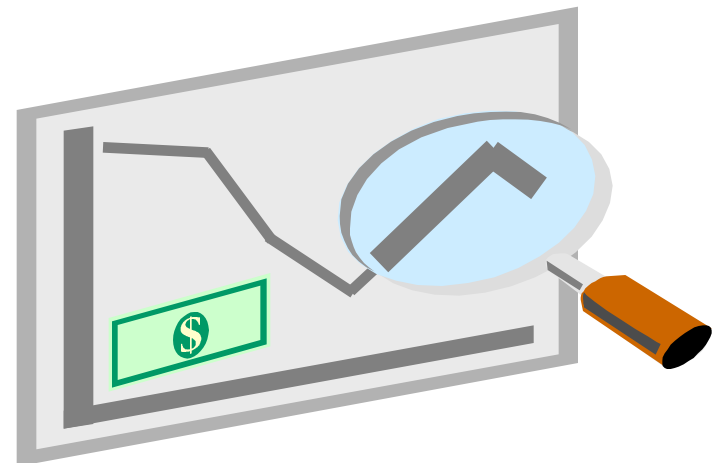
- **Motivacioni faktori poslovanja**

- unapređenje znanja o poslovanju
- odgovor izazovima i promenama tržišta
- investiranje u proizvod
- investiranje u tehnologije
- investiranje u kupce i zadržavanje kupaca
- obezbeđenje pristupa informacijama o poslovanju
- obezbeđenje profitabilnog poslovanja
- obezbeđenje konkurentnosti proizvoda



# Motivacija nastanka DW sistema

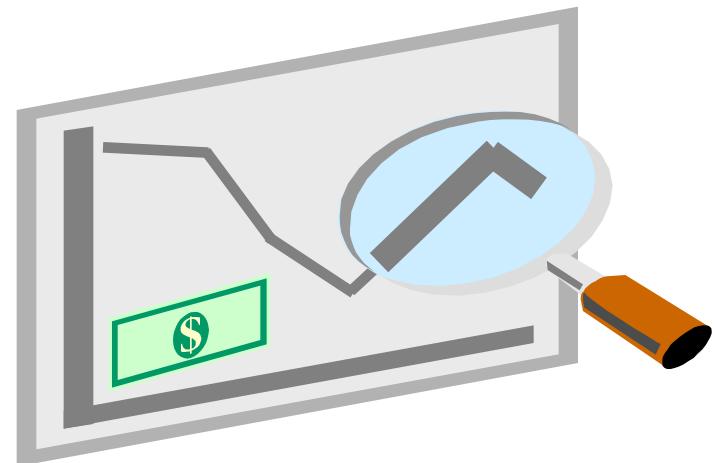
- **Uloga IS u podršci motivacionih faktora**
  - pružanje kvalitetnih informacija o poslovanju
    - smanjenje troškova poslovanja
    - praćenje i analiza poslovanja
    - povećanje profita
  - podrška upravljanju
    - podrška odlučivanju





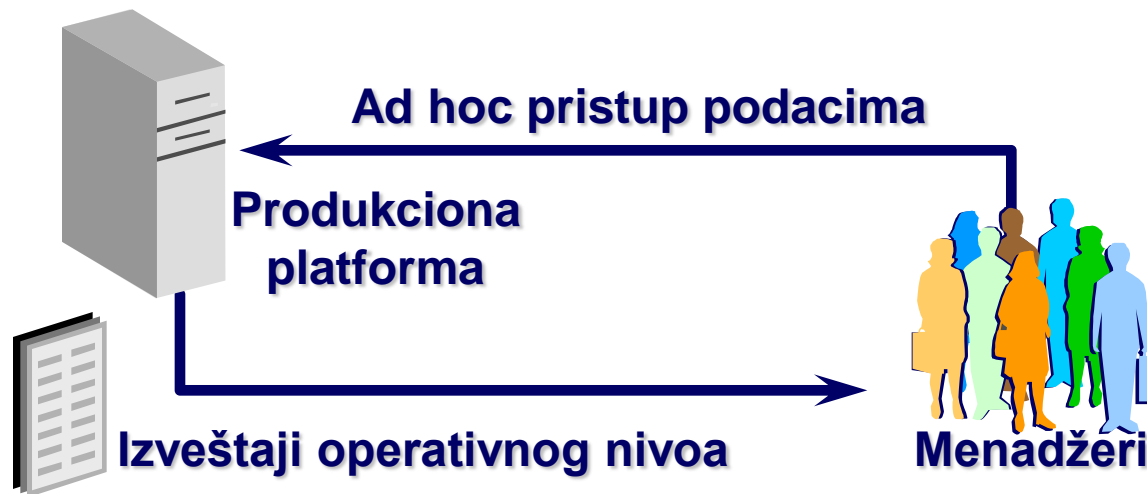
# Motivacija nastanka DW sistema

- **Podrška poslovanju**
  - **Upravljački informacijski sistemi**
    - Management Information Systems (MIS)
      - >1970 (>1980) godine
  - **Sistemi za podršku odlučivanju**
    - Decision Support Systems (DSS)
      - >1980 (>1990) godine



# Motivacija nastanka DW sistema

- **Upravljački informacijski sistemi**
  - obezbeđuju **operativne podatke** o poslovanju
  - **ad hoc pristup** operativnim podacima
  - izveštaji se kreiraju po zahtevima korisnika
  - skromne mogućnosti izveštaja u **podršci analize**



# Motivacija nastanka DW sistema

---

- **Upravljački informacijski sistemi**
  - sistemi za podršku operativnog, transakciono orijentisanog poslovanja
  - **On-line Transaction Processing (OLTP) sistemi**
    - projektovani da obezbede odlične performanse (kratko vreme odziva sistema) u realizaciji transakcija
    - kompleksne strukture podataka
    - nedovoljno dobro iskazana semantika struktura podataka
    - visoka disperzija podataka
    - nepogodni za realizaciju zahtevnih upita (upita nad velikim količinama podataka)

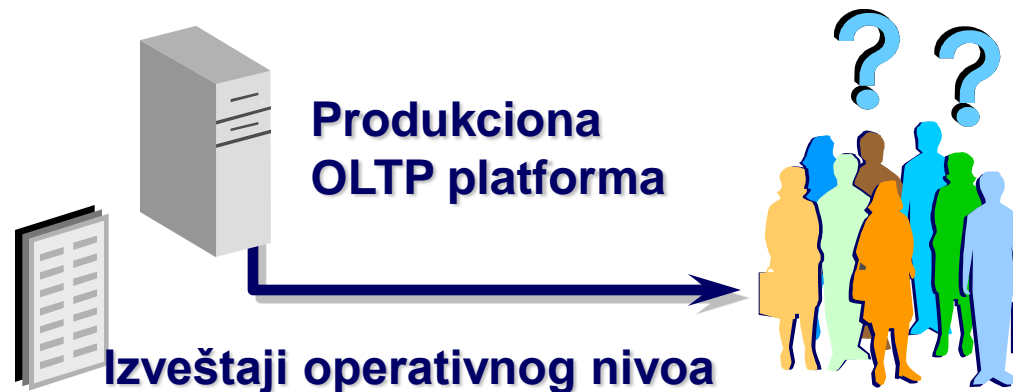
# Motivacija nastanka DW sistema

- Tehničke karakteristike OLTP sistema**

Karakteristika	Opis
Tipične operacije	Operacije ažuriranja
Količina podataka po transakciji	Mala / Srednja
Jedinica obrade podataka	Slog (zapis, n-torka)
Struktura formi za prikaz podataka	Statička, retko promenljiva
Nivo agregiranosti podataka	Nizak
Vremensko određenje podataka	Trenutno stanje (dan)
Nivo podrške analize podataka	Nizak

# Motivacija nastanka DW sistema

- **Upravljački informacijski sistemi**
  - analiza podataka, korišćenjem OLTP sistema – otežana i nepogodna
    - svodi se na
      - analizu operativnih podataka
      - korišćenje izveštaja koji se direktno generišu nad operativnim podacima



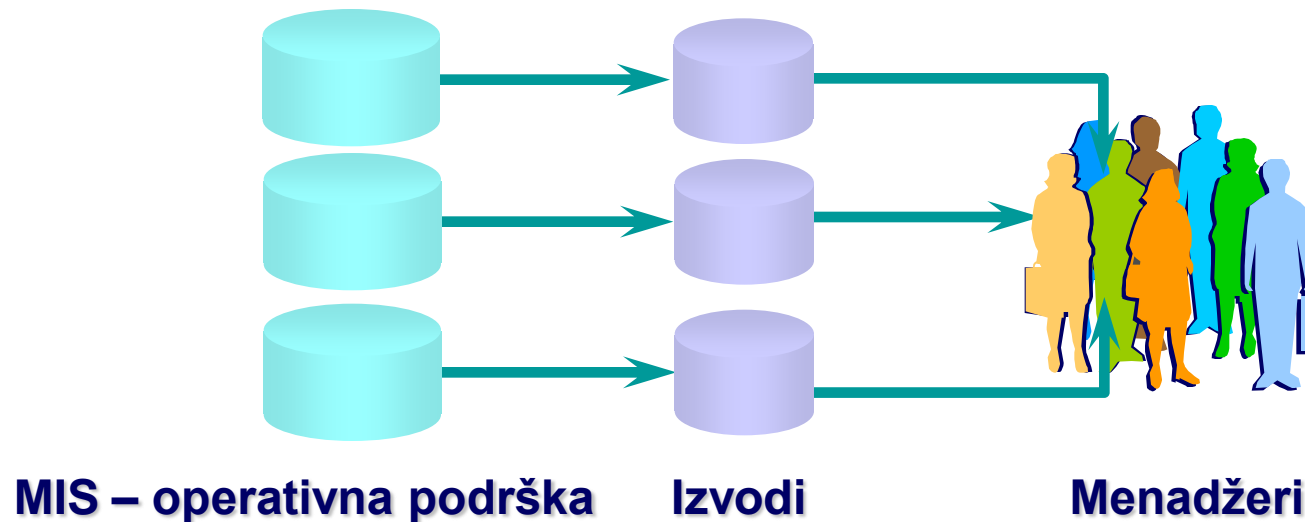
# Motivacija nastanka DW sistema

- **OLTP / Proces analize podataka**
  - Mogućnosti OLTP / Potrebe kompleksnih analiza

<b>OLTP</b>	<b>Kompleksne analize</b>
Informaciona podrška dnevnog poslovanja	<b>Analiza istorijskih podataka</b>
Skladištenje podataka na nivou pojedinačnih transakcija	<b>Analiza agregiranih (integrisanih) podataka</b>
Šema baze podataka: normalizovana	<b>Šema baze podataka: de-normalizovana (zvezdasta)</b>

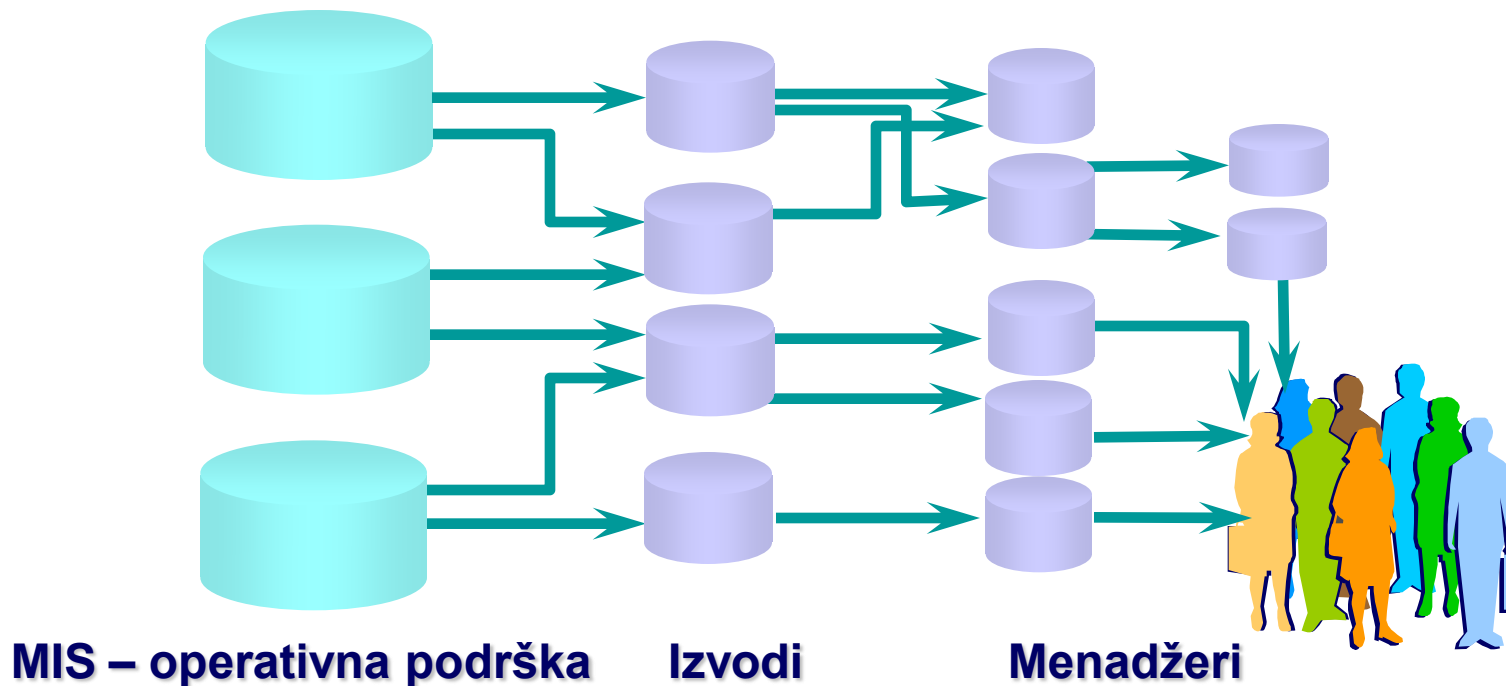
# Motivacija nastanka DW sistema

- **MIS u funkciji obezbeđenja analiza**
  - "primitivno" rešenje
    - obrada posebno izdvojenih (ekstrahovanih) podataka
      - postaje **nezavisna** (odvaja se) od samog MIS-a
      - stvaraju se "**privatni**" podaci korisnika



# Motivacija nastanka DW sistema

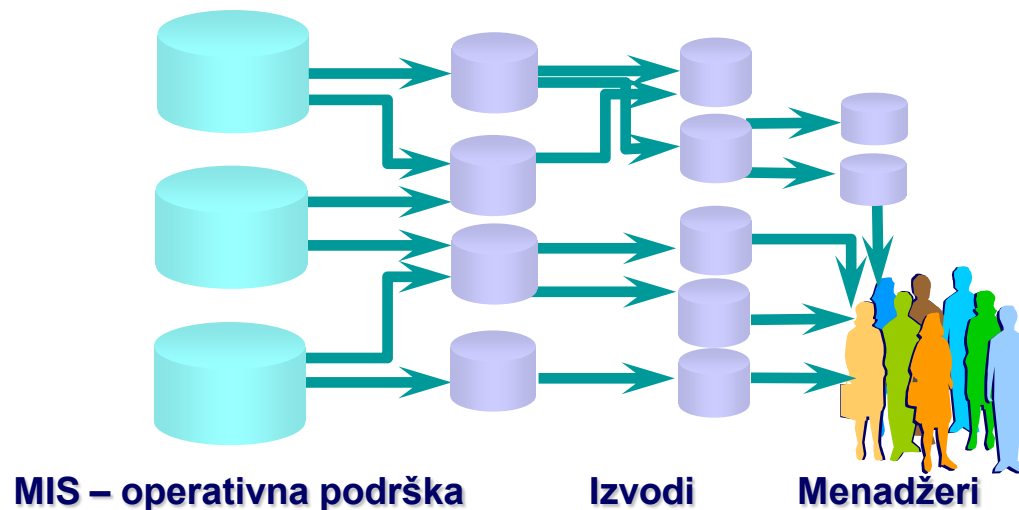
- **Obrada posebnih izvoda podataka**
  - problemi
    - upravljački aspekt
      - **"eksplozija" izvoda podataka**





# Motivacija nastanka DW sistema

- **Obrada posebnih izvoda podataka**
  - problemi
    - aspekt efikasnosti (produktivnosti) poslovanja
      - višestruki napor za rešavanje istih problema
      - upotreba različitih tehnologija
      - pojava nepotrebnih izveštaja
      - nepostojanje meta podataka



# Motivacija nastanka DW sistema

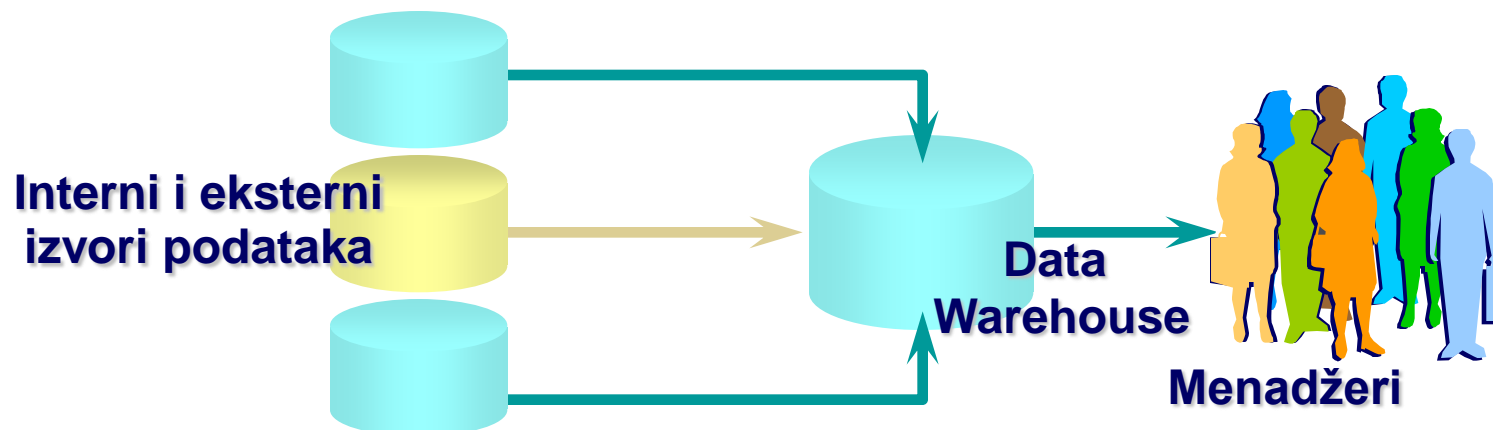
---

- **Obrada posebnih izvoda podataka**
  - problemi
    - **aspekt podataka**
      - **ne postoji opšta vremenska odrednica podataka**
      - **primena različitih algoritama obrade za iste probleme**
      - **izdvajanje podataka na različitim nivoima**
      - **različit nivo granularnosti podataka**
      - **neusaglašenost naziva i značenja atributa**
      - **gubitak informacija**
      - **nepostojanje procedura za usaglašavanje podataka**
      - **nepostojanje "drill-down" izveštajnih mehanizama**



# Motivacija nastanka DW sistema

- **DSS u funkciji obezbeđenja analiza**
  - savremeno i bolje rešenje
    - softverska podrška postupaka analize podataka
      - u funkciji **podrške odlučivanja i upravljanja**
      - **uvažava potrebe i mogućnosti poslovnog sistema u celini**
      - zasnovana na **jedinstvenom "skladištu" podataka**



# Motivacija nastanka DW sistema

---

- **Sistem za podršku odlučivanju (DSS)**
  - programski sistem namenjen za podršku rukovodiocima pri donošenju poslovnih odluka
    - podržava proces odlučivanja
    - u cilju postizanja uspeha preduzeća na tržištu
  - koristi se u procesu odlučivanja na svim nivoima organizacione strukture



# Motivacija nastanka DW sistema

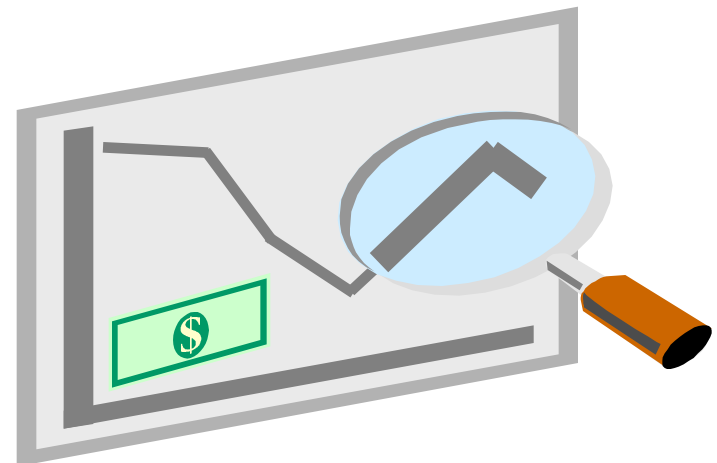
---

- **Sistem za podršku odlučivanju (DSS)**
  - ilustracija načina upotrebe DSS sistema
    - davanje odgovora na pitanja tipa
      - Kakav je odnos rasta produktivnosti u posmatranim organizacionim jedinicama preduzeća za period od poslednje tri godine?
      - Kakav je odnos uložениh sredstava u reklamiranje proizvoda i ostvarenih sredstava od njegove prodaje, u poslednja tri meseca?



# Motivacija nastanka DW sistema

- **Sistem za podršku odlučivanju (DSS)**
  - koristi istorijske operativne podatke kao ulaz
  - podržava matematičke modele analize poslovanja
  - produkuje informacije, bitne za proces odlučivanja



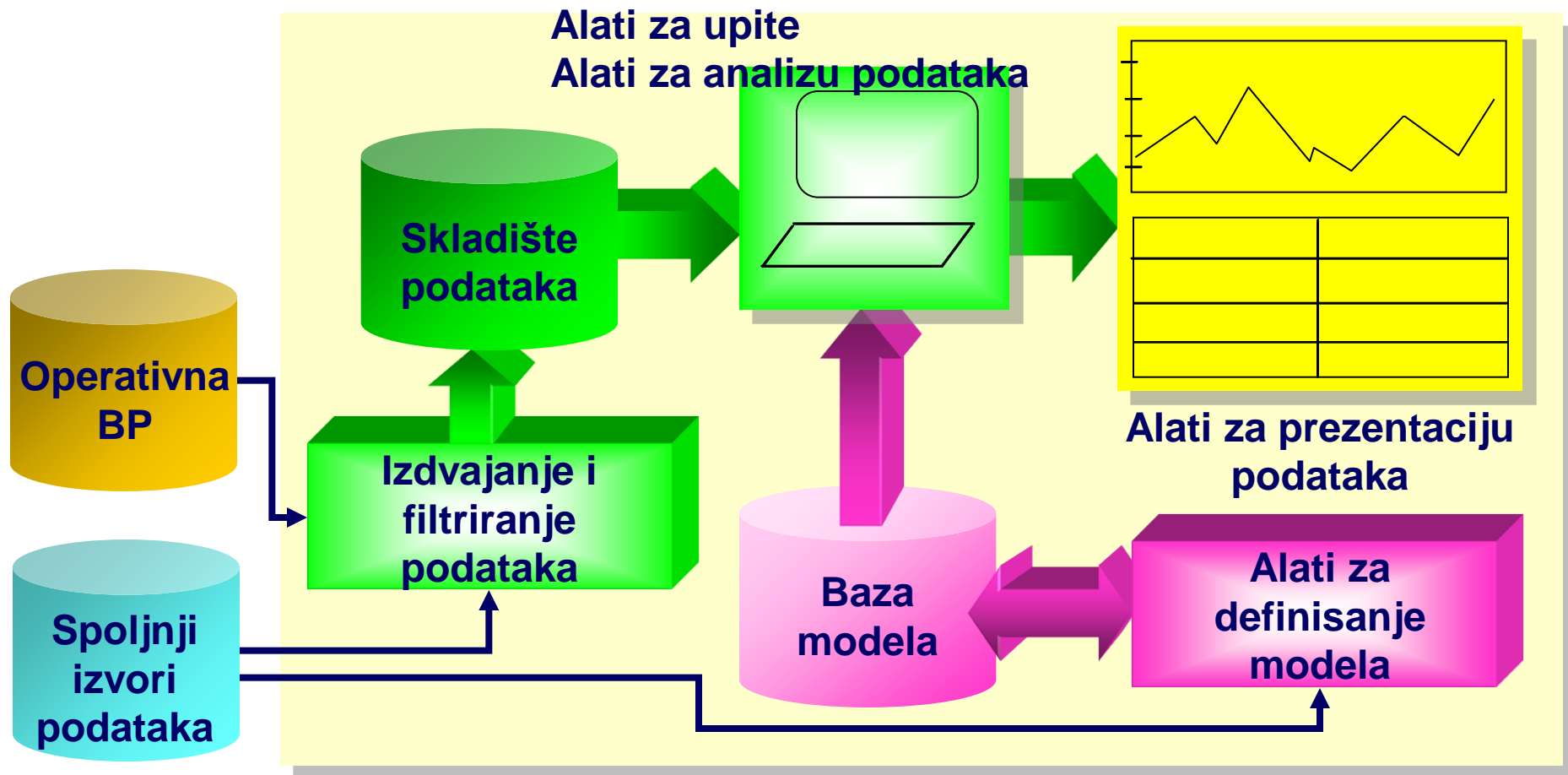
# Motivacija nastanka DW sistema

---

- **Opšta struktura DSS sistema**
  - **Komponenta za upravljanje podacima**
    - skladište podataka (Data Warehouse)
    - alati za ekstrahovanje i filtriranje podataka
    - alati za upite
  - **Komponenta za upravljanje modelima**
    - baza modela za analizu i odlučivanje
    - alati za definisanje modela analize i odlučivanja
    - alati za izvršavanje modela analize i odlučivanja
  - **Komponenta za prezentaciju podataka**

# Motivacija nastanka DW sistema

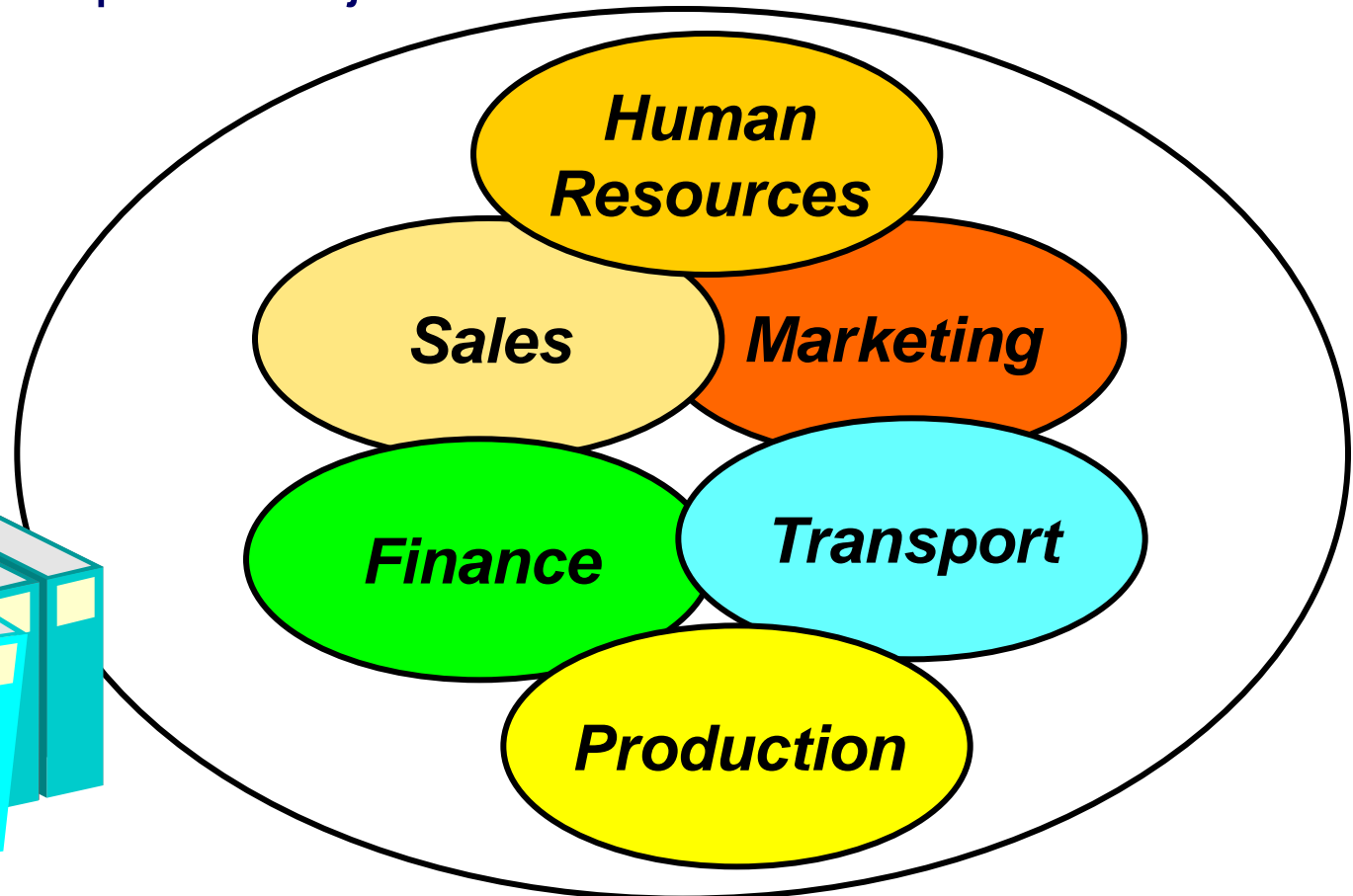
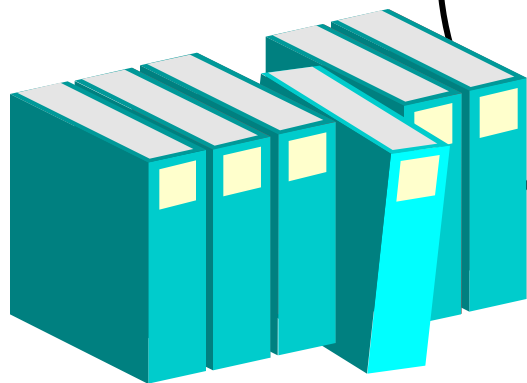
- Opšta struktura DSS sistema





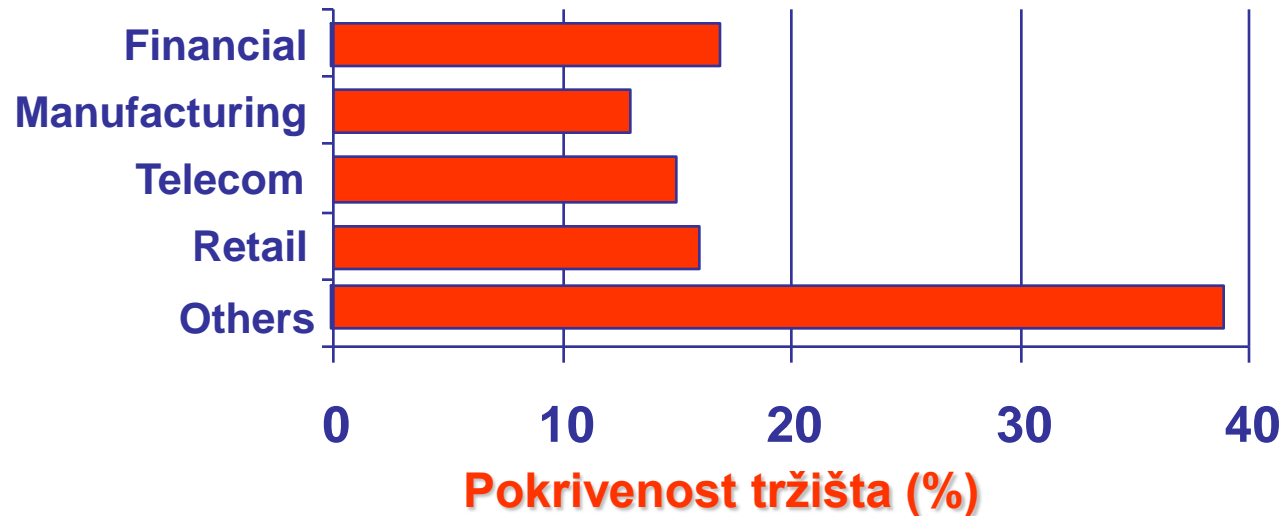
# Motivacija nastanka DW sistema

- **Oblasti primene DSS & DW sistema**
  - svi aspekti poslovanja



# Motivacija nastanka DW sistema

- **Delatnosti za primenu DSS & DW sistema**



- Airline
- Banking
- Health care
- Investment
- Insurance
- Retail
- Telecommunications
- Manufacturing
- Credit card suppliers
- Clothing distributors

# Motivacija nastanka DW sistema

---

- **Tehnološki motivacioni faktori za primenu DSS & DW sistema**
  - **računarsko-komunikacione tehnologije**
    - 64-bitni procesori, RISC, SMP, Cluster, MPP, NUMA arhitekture
    - RAID jedinice diskova
    - >Gb propusna moć mreže
  - **operativni sistemi**
    - Mrežne, Cluster i Grid arhitekture

# Motivacija nastanka DW sistema

---

- **Tehnološki motivacioni faktori za primenu DSS & DW sistema**
  - **systemi za upravljanje bazama podataka**
    - paralelno izvršavanje upita
    - fizičko particioniranje baze podataka
    - tehnike indeksiranja
    - tehnike distribucije i replikacije BP
  - **višeslojne K/S arhitekture sistema**
  - **namenski alati za izgradnju DSS & DW sistema**
  - **web orijentisane softverske tehnologije**
  - **Internet**

# Sadržaj

---

- Motivacija nastanka DW sistema
- **Koncepcija DW sistema**
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Koncepcija DW sistema

---

- **Data Warehouse (DW)**
  - pojam nastaje kao odgovor na zahtev da
    - DSS treba da poseduju i obrađuju veliku količinu podataka, proizvedenih iz operativnih podataka
    - je potreban **novi pristup organizaciji i upravljanju** podacima za DSS
  - DW predstavlja osnovni element DSS nove generacije
  - pojam "**skladište podataka**" ne iskazuje, na **zadovoljavajući način**, suštinu pojma DW
    - to nije baza operativnih podataka,
    - niti nestrukturirani skup podataka

# Koncepcija DW sistema

- **Data Warehouse (DW)**

"A Data Warehouse is a

- subject oriented,
- integrated,
- time variant,
- non volatile

collection of data in support of management's decision making process."

- **Bill Inmon**
  - tvorac pojma DW

# Koncepcija DW sistema

---

- **Data Warehouse (DW)**

"A Data Warehouse is an enterprise structured repository of

- subject oriented,
- time variant,
- historical

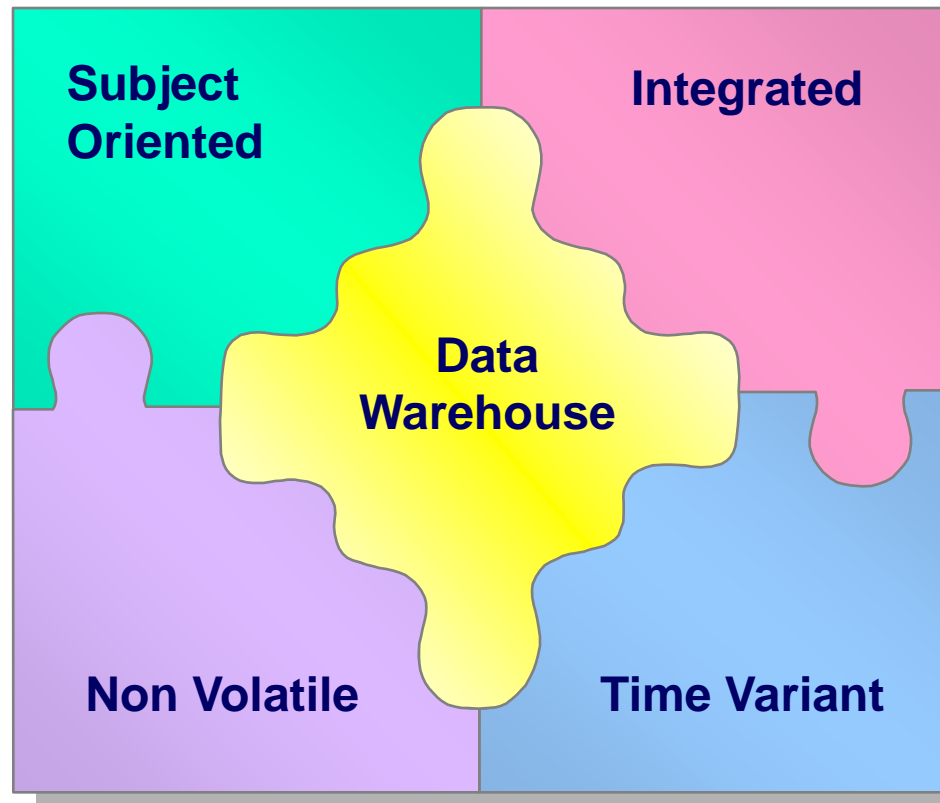
data used for information retrieval and decision support. The data warehouse stores atomic and summary data."

- Oracle Data Warehouse Method



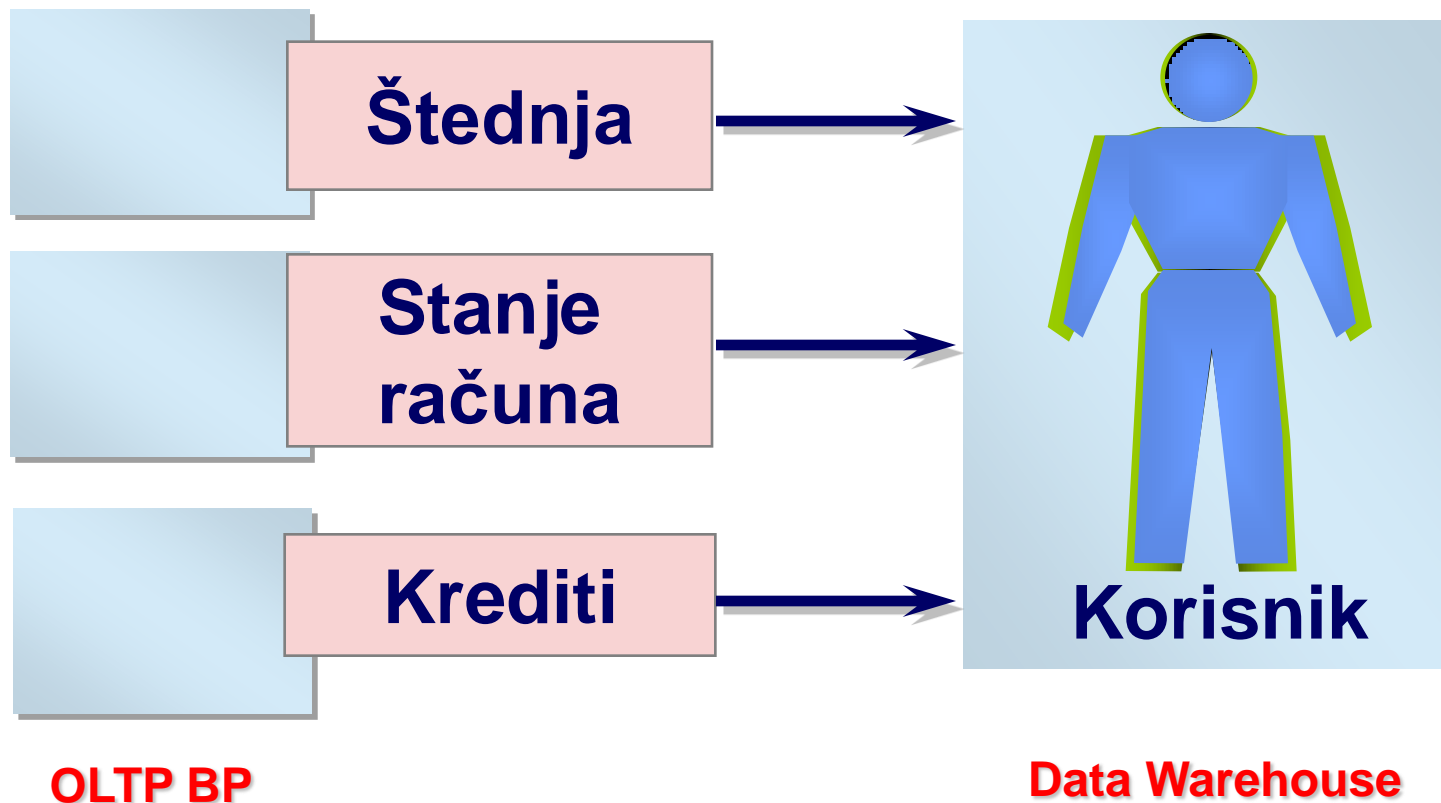
# Koncepcija DW sistema

- Opšte odrednice DW sistema



# Integrisanost DW

- Podaci o jednom entitetu se zadaju i čuvaju na jednom mestu



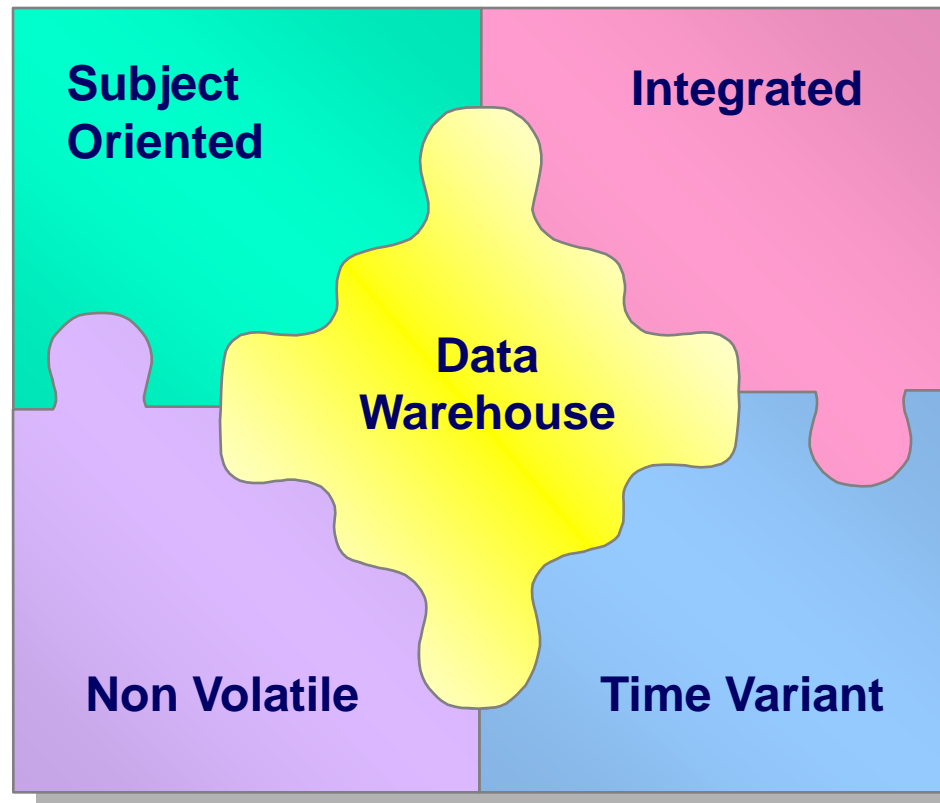
# Integriranost DW

---

- Data Warehouse predstavlja **centralizovanu bazu podataka**
  - sadrži podatke svih organizacionih delova preduzeća
  - u "standardizovanom" formatu
    - često, podaci iz operativnih BP su particionirani u više organizacionih delova
    - pojedinačni delovi BP su različito strukturirani
      - mogu biti implementirani čak i pod različitim SUBP-ovima

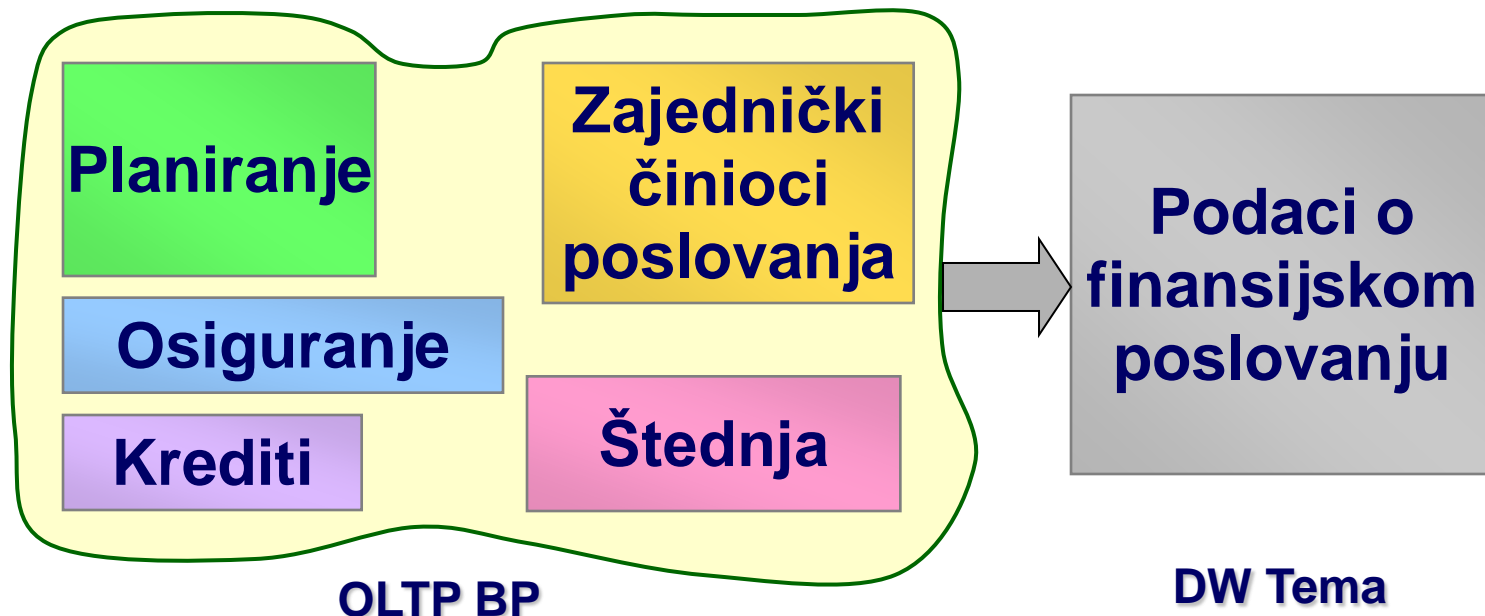
# Koncepcija DW sistema

- Opšte odrednice DW sistema



# Orientacija DW sistema na teme

- Podaci se **kategorišu i organizuju po temama poslovanja**
  - a ne po funkcionalnim celinama, kao u OLTP sistemima



# Orijentacija DW sistema na teme

---

- Posledica zahteva da
  - različite funkcije poslovnog sistema treba da upotrebljavaju **jedinstveni Data Warehouse**
- odnos teme – funkcije poslovnog sistema
  - jedna tema može biti od interesa za više poslovnih funkcija, i obratno
  - različite teme mogu deliti zajedničke podatke

# Orientacija DW sistema na teme

---

- **Primeri mogućih tema**

- prodaja

- podaci o **proizvodima**, kupcima, prostornoj i organizacionoj strukturi,...

- marketing

- podaci o tržištu, **proizvodima**, kupcima, konkurenciji,...

- proizvodnja

- podaci o **proizvodima**, kupcima, tehnologijama, planovima,...

- finansijsko poslovanje

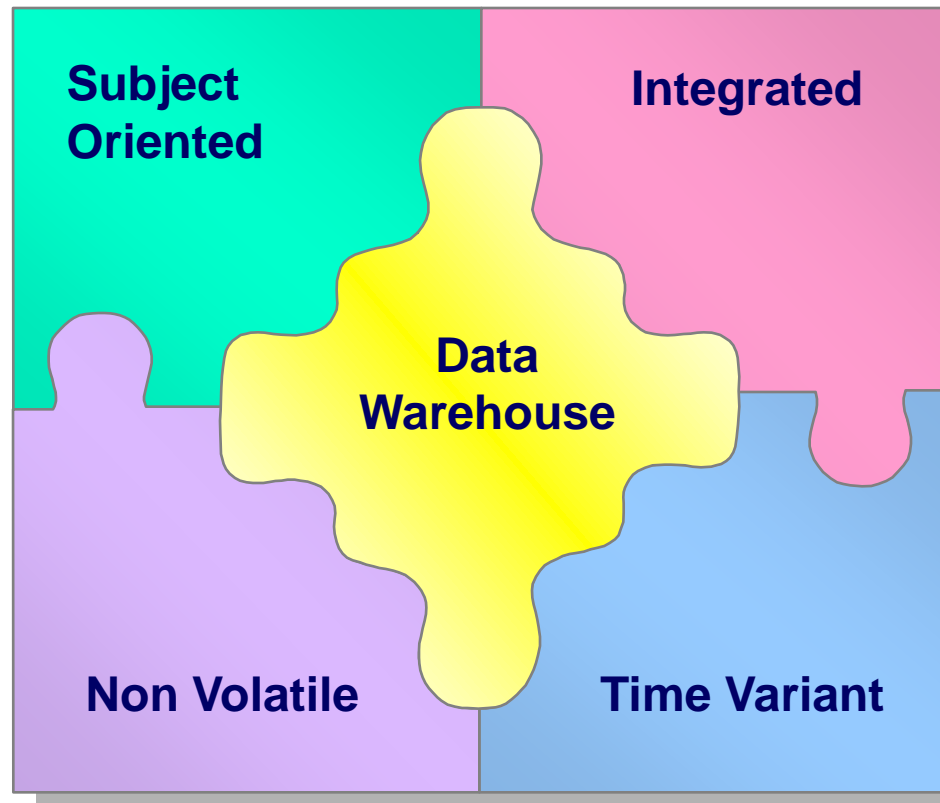
- podaci o **proizvodima**, kupcima, vrstama i dinamici troškova,...

- transport

- ...

# Koncepcija DW sistema

- Opšte odrednice DW sistema

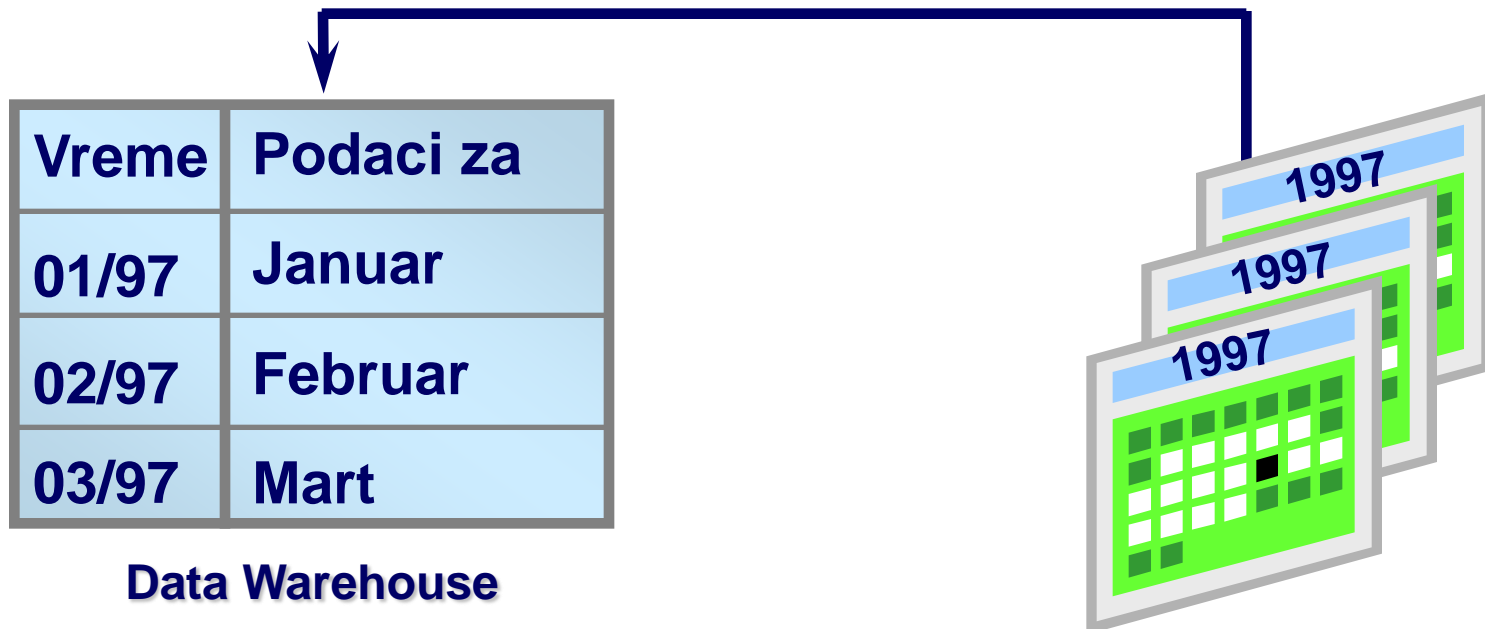




# Vremenska zavisnost DW

- **Zavisnost podataka u DW**

- podaci se organizuju putem nizova **"preseka stanja"** ("slika") operativnih podataka
  - svaki presek se odnosi na **jedan vremenski interval**



# Vremenska zavisnost DW

---

- **Zavisnost podataka u DW**
  - DW **sadrži vreme, kao dimenziju**
    - DW sadrži agregirane podatke o poslovanju u vremenu
      - **istorijske**
        - » poslednja nedelja, mesec, godina, obračunski period
      - **u budućnosti**
        - » podaci koji reprezentuju **predviđanja**, kao rezultat sprovedenih analiza

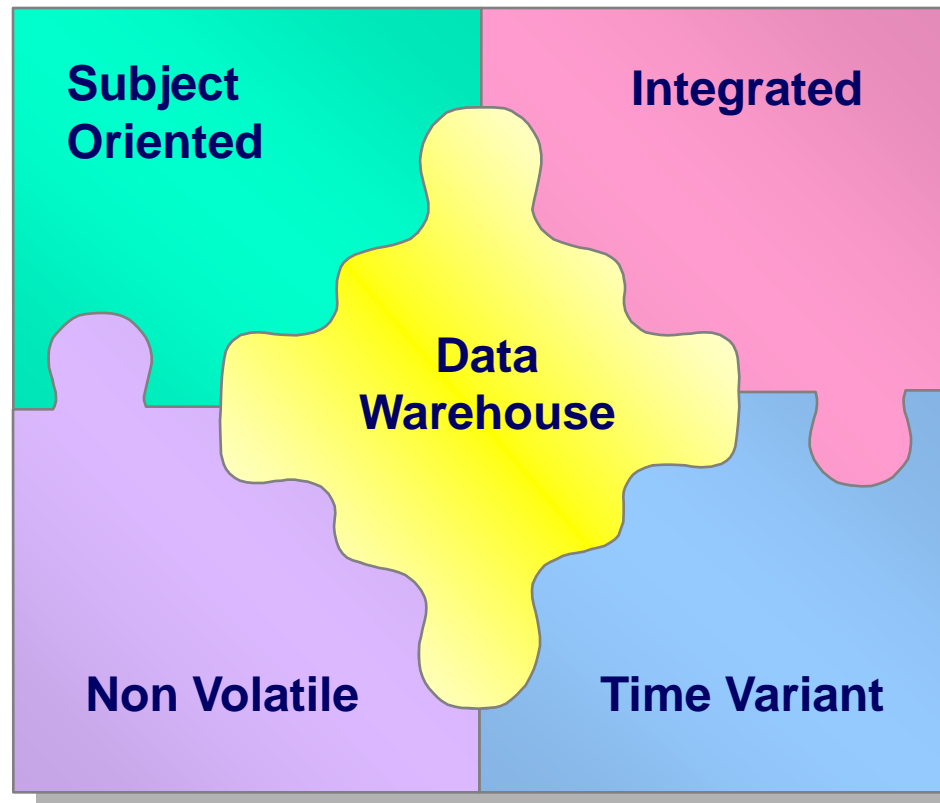
# Vremenska zavisnost DW

---

- **Zavisnost DW od vremena**
  - DW se **periodično ažurira**
    - dodavanjem novih podataka, preuzetih iz operativnih BP i ostalih spoljnjih izvora
    - ponovnim izračunavanjem vrednosti vremenski zavisnih agregiranih podataka, kao posledica dodavanja podataka
  - Primer
    - ukoliko se u DW dodaju najnoviji podaci o prodaji za poslednji vremenski period, potrebno je preračunati sumarne podatke o prodaji za proizvode, kupce i prodajna mesta

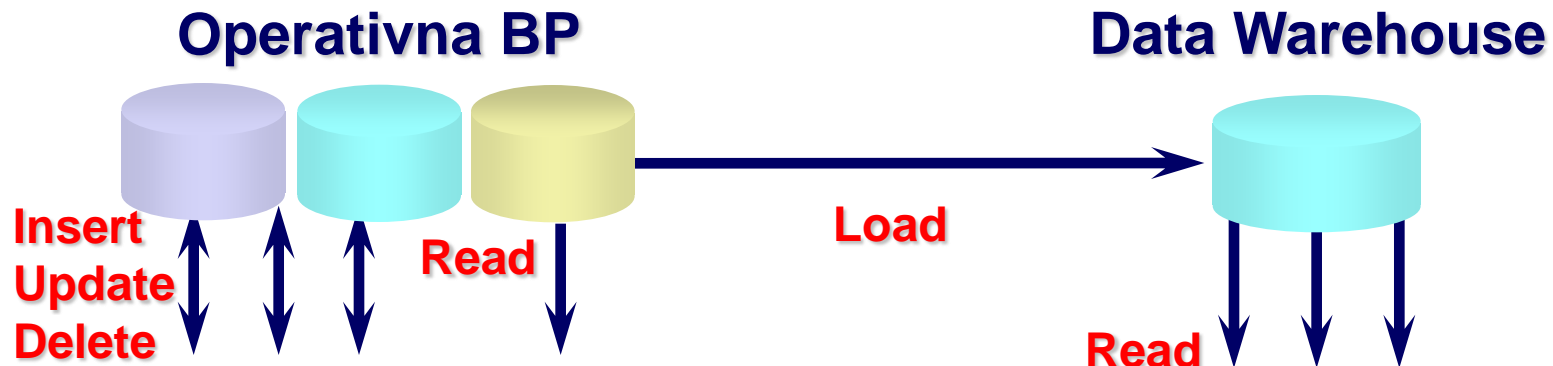
# Koncepcija DW sistema

- Opšte odrednice DW sistema



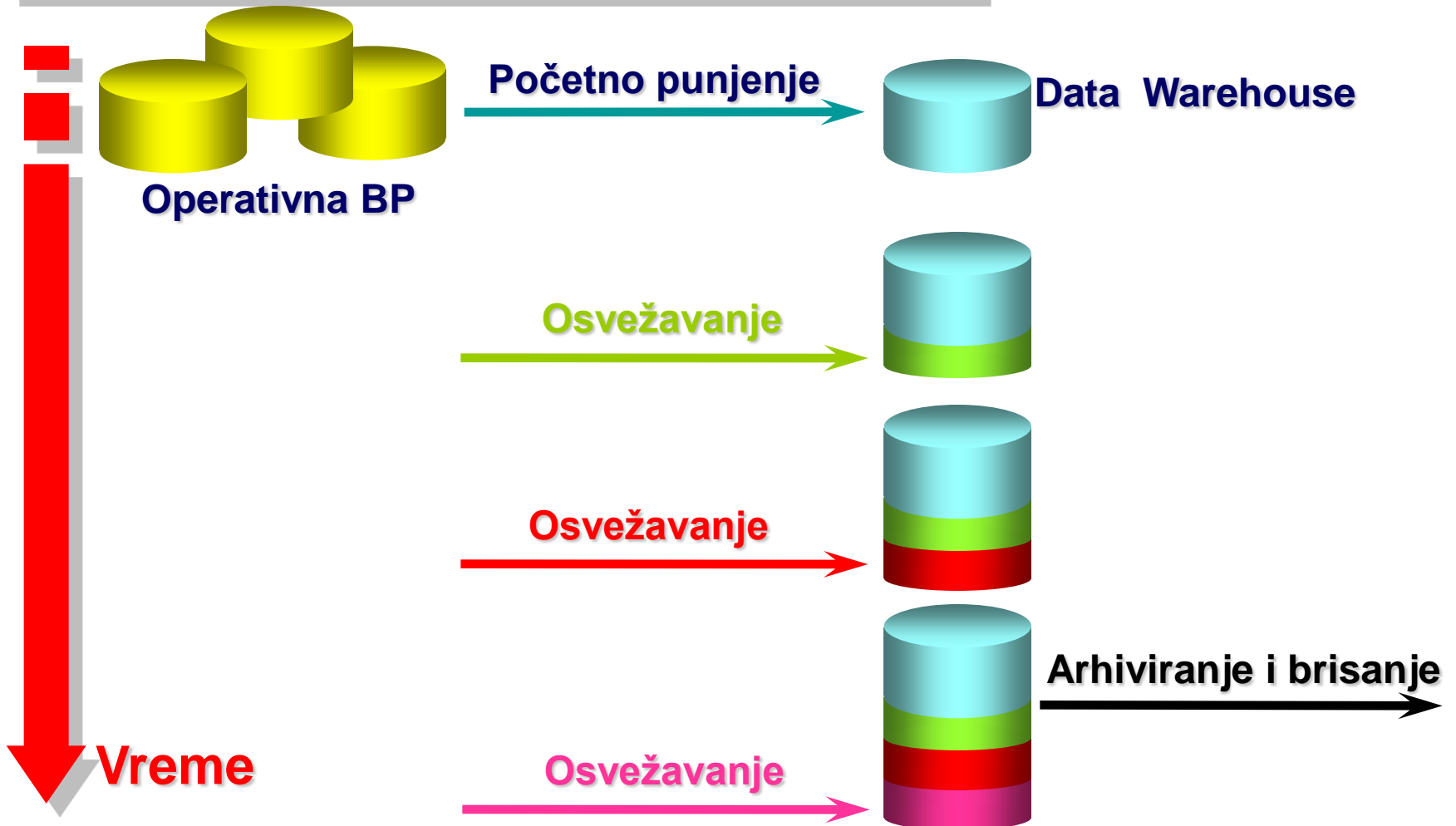
# Nepromenljivost DW

- Podaci u DW se, praktično, **vrlo retko brišu ili modifikuju**
  - Najčešće, **podaci se samo dodaju** u DW
  - DW se uvek **samo povećava u vremenu**
  - dostiže **razmere TByte baza podataka**
    - SUBP mora biti osposobljen da upravlja, po obimu podataka, vrlo velikim bazama podataka



# Nepromenljivost DW

## Opšti postupak ažuriranja DW



# Sadržaj

---

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Tematske karakteristike DW

---

- Dimenzionost
- Granularnost
- Diskretizacija vremena



# Dimenzionost

---

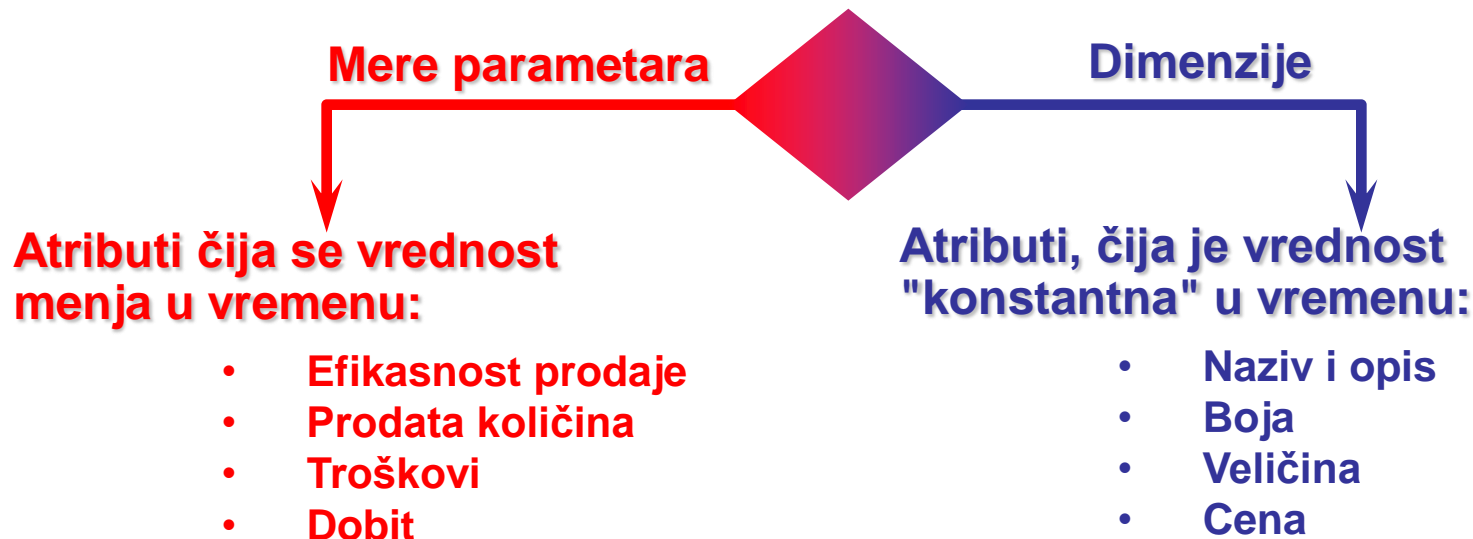
- **Dimenzionalna (višedimenzionalna)** organizovanost podataka u DW
  - DW sadrži podatke o vrednostima praćenih parametara poslovanja
    - takvi podaci uvek se iskazuju s obzirom na različite kontekste (dimenzije)
  - Vrste podataka u DW
    - **činjenični podaci**
      - iskazuju vrednosti praćenih parametara poslovanja, u okviru zadatog konteksta (dimenzionih parametara)
    - **dimenzioni podaci**
      - iskazuju dimenzione karakteristike činjeničnih podataka

# Dimenzionost

- **Višedimenzionalnost DW**

- Vrste podataka u DW

- činjenični podaci
- dimenzioni podaci

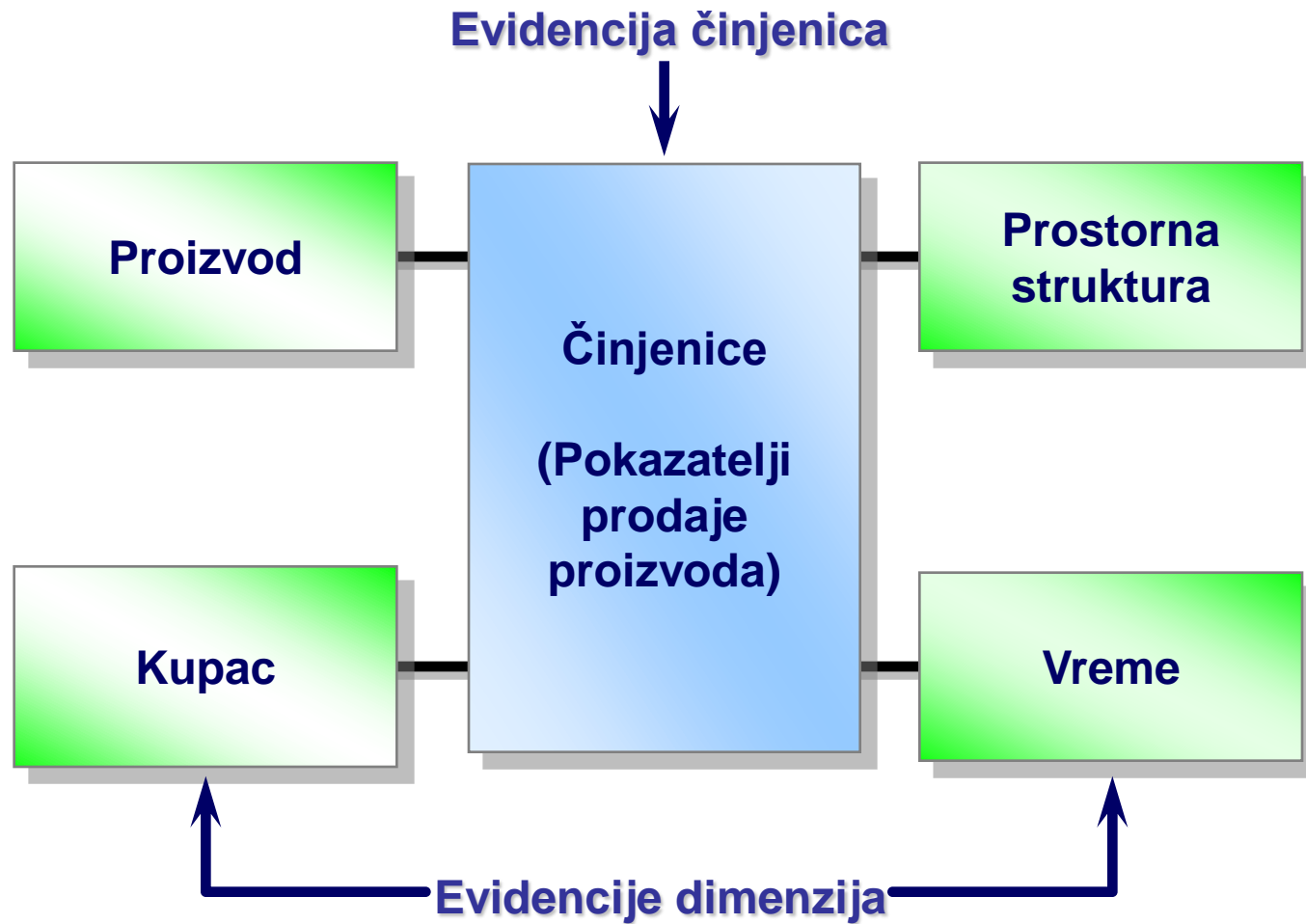


# Dimenzionost

---

- **Motivacija višedimenzionalnosti DW**
  - donošenje poslovnih odluka zahteva međusobno povezivanje kako pojedinačnih, tako i agregiranih podataka o entitetima različitih klasa
  - Primer
    - Praćenje i analiza prodaje proizvoda
      - **dimenzije**
        - » kupac, proizvod, vreme, prostorna struktura
      - **činjenice**
        - » prodana količina posmatranog proizvoda po kupcu, u zadatom vremenskom intervalu i u zadatom prostoru
        - » bruto promet, dobit od prodaje, troškovi prodaje,...

# Dimenzionost



# Tematske karakteristike DW

---

- Dimenzionost
- Granularnost
- Diskretizacija vremena

# Granularnost

---

- Operativni podaci su predmet individualnih transakcija u OLTP sistemu
- Podaci za DW se dobijaju **agregacijom (sintetizovanjem)** operativnih podataka
  - nivoi, na kojima se agregiraju podaci, mogu biti **različito definisani**
    - u zavisnosti od definisanih poslovnih ciljeva

# Granularnost

---

- **Granularnost**

- izbor osnovne jedinice mere posmatrane dimenzije

- **fina granulacija**

- viši stepen granularnosti
    - najviši stepen granularnosti (ekstremni slučaj)
      - operativni podaci iz OLTP sistema

- **gruba granulacija**

- niži stepen granularnosti
    - podaci iz DW

- **nivo granularnosti**

- definisan nivoom agregacije podataka u DW
      - obrnuto proporcionalan nivou agregacije podataka

**Primer  
vremenski aspekt  
granularnosti DW**

# Granularnost

**Primer  
prostorni aspekt  
granularnosti DW**

## Nivo agregacije podataka u DW

- definisan "najužim" kontekstom iskazivanja podataka u DW – najmanjom jedinicom mere





# Granularnost

---

- **Nivo agregacije podataka u DW**

- Primer

- menadžer analizira podatke o prodaji s obzirom na prostorni (geografski) aspekt
      - Potrebni su mu podaci na nivou
        - » **prodajnog mesta,**
        - » **grada i**
        - » **regiona**
      - Nisu mu nikada potrebni podaci o **pojedinačnim transakcijama prodaje**

# Tematske karakteristike DW

---

- Dimenzionost
- Granularnost
- Diskretizacija vremena

# Diskretizacija vremena

---

- **Vreme u OLTP sistemima**

- operativni podaci u BP imaju **kratak životni vek**

- u principu, do jedne godine
- odnose se na pojedinačne poslovne transakcije
- relativno brzo postaju nepotrebni i brišu se (arhiviraju)
- vreme diskretizovano do jedinice u rasponu
  - dan (npr. bankarstvo, osiguranje, generalno obračunska funkcija u organizacijama), do
  - sekunda (telekom usluge, “billing systems”)
- nepogodni su za odlučivanje strateškog ili taktičkog karaktera

# Diskretizacija vremena

---

- **Vreme u DW sistemima**
  - praktično, obavezna dimenzija podataka u DW
  - DW obuhvata relativno **dugačak period vremena**
    - u principu nekoliko godina (i do 10 godina)
    - podaci se odnose na parametre poslovanja
    - vreme diskretizovano do jedinice u rasponu
      - dan, do
      - mesec

# Diskretizacija vremena

---

- **Vreme u DW sistemima**
  - izbor vremenske jedinice diskretizacije
    - prema potrebama menadžmenta strateškog i taktičkog nivoa
    - svi agregirani podaci se iskazuju s obzirom na
      - zadatu vremensku jedinicu diskretizacije, ili
      - sumiraju se na višu vremensku jedinicu diskretizacije
  - **kritični faktor uspeha** u životnom veku DW sistema

# Sadržaj

---

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema



# Poređenje karakteristika DW i OLTP

Transakcione karakteristike

Karakteristika	Operativna BP (OLTP)	Data Warehouse
Tipične operacije	Operacije ažuriranja	Operacije upita
Kritične transakcije	Transakcije ažuriranja	Transakcije upita
Ažuriranje BP	Veliki broj DML operacija	Punjenje i periodično osvežavanje
Frekvencija upita	Niska / Srednja	Visoka
Kompleksnost upita	Niska	Visoka
Količina podataka po transakciji	Mala / Srednja	Velika
Očekivano vreme odgovora	Do nivoa sekunde	Nekoliko sekundi – Više sati

# Poređenje karakteristika DW i OLTP

Karakteristike podataka

Karakteristika	Operativna BP (OLTP)	Data Warehouse
Vremenska diskretizacija podataka	Dan – Sekunda	Dan – Godina
Aktuelnost podataka	Do jedne godine	Više godina
Obim baze podataka	MB – GB	GB – TB
Povećanje obima baze podataka	Linearno	Polinomijalno / Eksponencijalno
Granularnost podataka	Elementarni podaci	Agregirani podaci
Nivo agregacije	Nizak	Visok
Šema BP	Normalizovana, kompleksnija	Denormalizovana, manje kompleksna



# Poređenje karakteristika DW i OLTP

Karakteristike poslovanja

Karakteristika	Operativna BP (OLTP)	Data Warehouse
Izvori podataka	Operativno poslovanje	Operativna BP, interni i eksterni izvori
Organizacija podataka	Prema funkcijama	Prema temama
Podrška poslovnih procesa	Operativno poslovanje	Analiza i odlučivanje
Forme za prikaz podataka	Statičke, retko promenljive	Kontekstno zavisne, promenljive
Intenzitet korišćenja BP u vremenu	Uniforman	Neuniforman, mogući "udarni" termini

# Sadržaj

---

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Arhitektura DW sistema

- **Arhitektura DW sistema**

- praktično višeslojna i obuhvata

- **servere DW**

- sa odgovarajućim SUBP

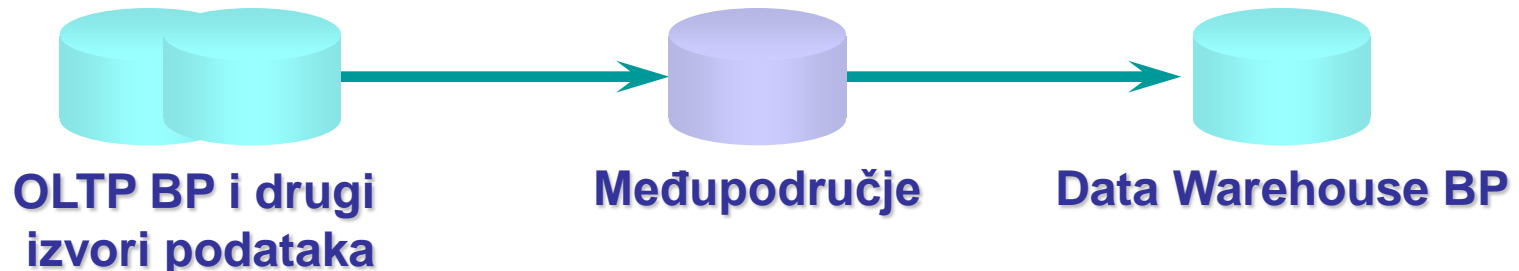
- **međupodručja DW sistema**

- **Data Staging Area (DSA), Operational Data Store (ODS)**

- memorišu podatke taktičkog nivoa iz operative BP ili iz drugih spoljnjih izvora

- mesto operative pripreme podataka za upis u DW sistem

- sa softverskom podrškom za kreiranje, osvežavanje i korišćenje DW baze podataka



# Arhitektura DW sistema

- **Arhitektura DW sistema**

- praktično višeslojna i obuhvata

- **DW aplikativne servere**

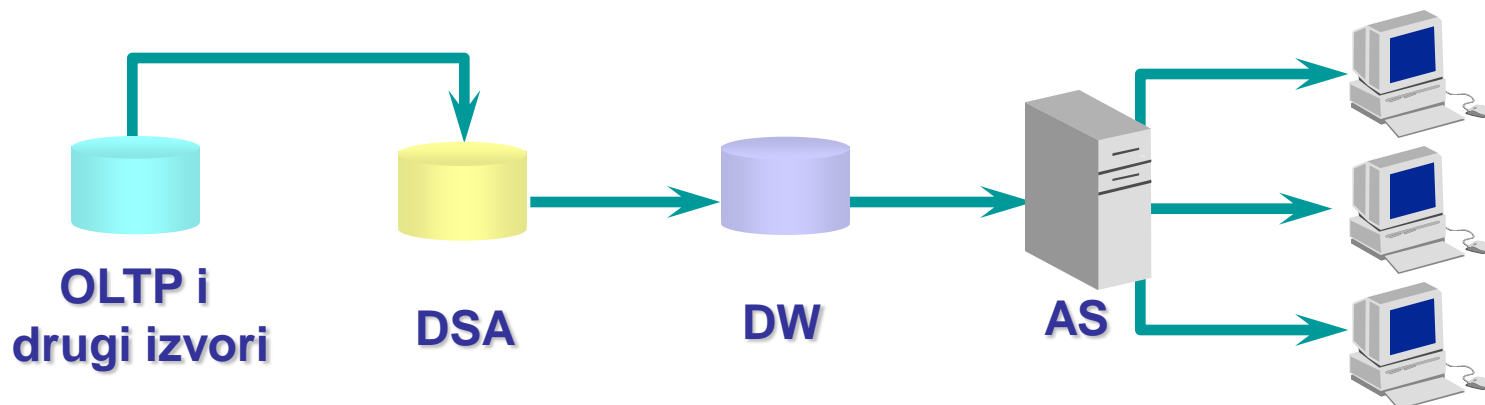
- sa aplikativnim softverom za podršku korišćenja DW baze podataka

- » upiti i prezentacija podataka

- » analiza, istraživanje podataka i proces odlučivanja

- **DW klijente**

- koji obezbeđuju UI za pristup DW sistemu



# Arhitektura DW sistema

- **Arhitektura DW sistema**

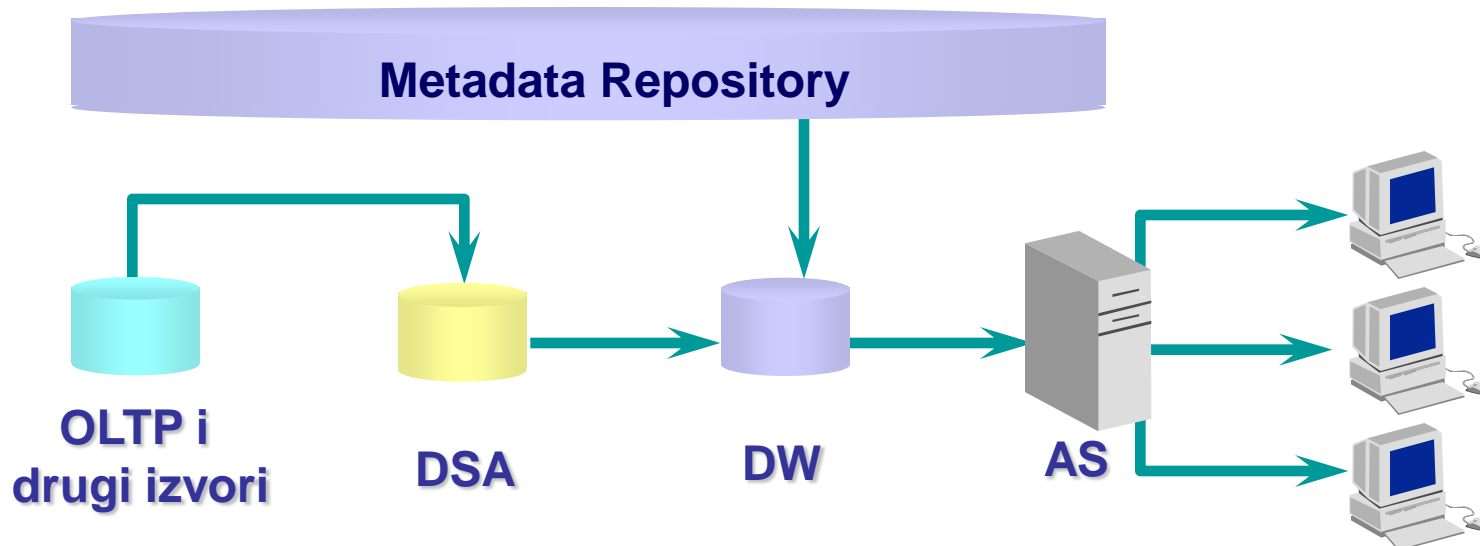
- praktično višeslojna i obuhvata

- **rezpozitorijum metapodataka**

- **Metadata Repository (MDR)**

- podatke o strukturi i sadržaju samog DW sistema

- » podatke o podacima u DW



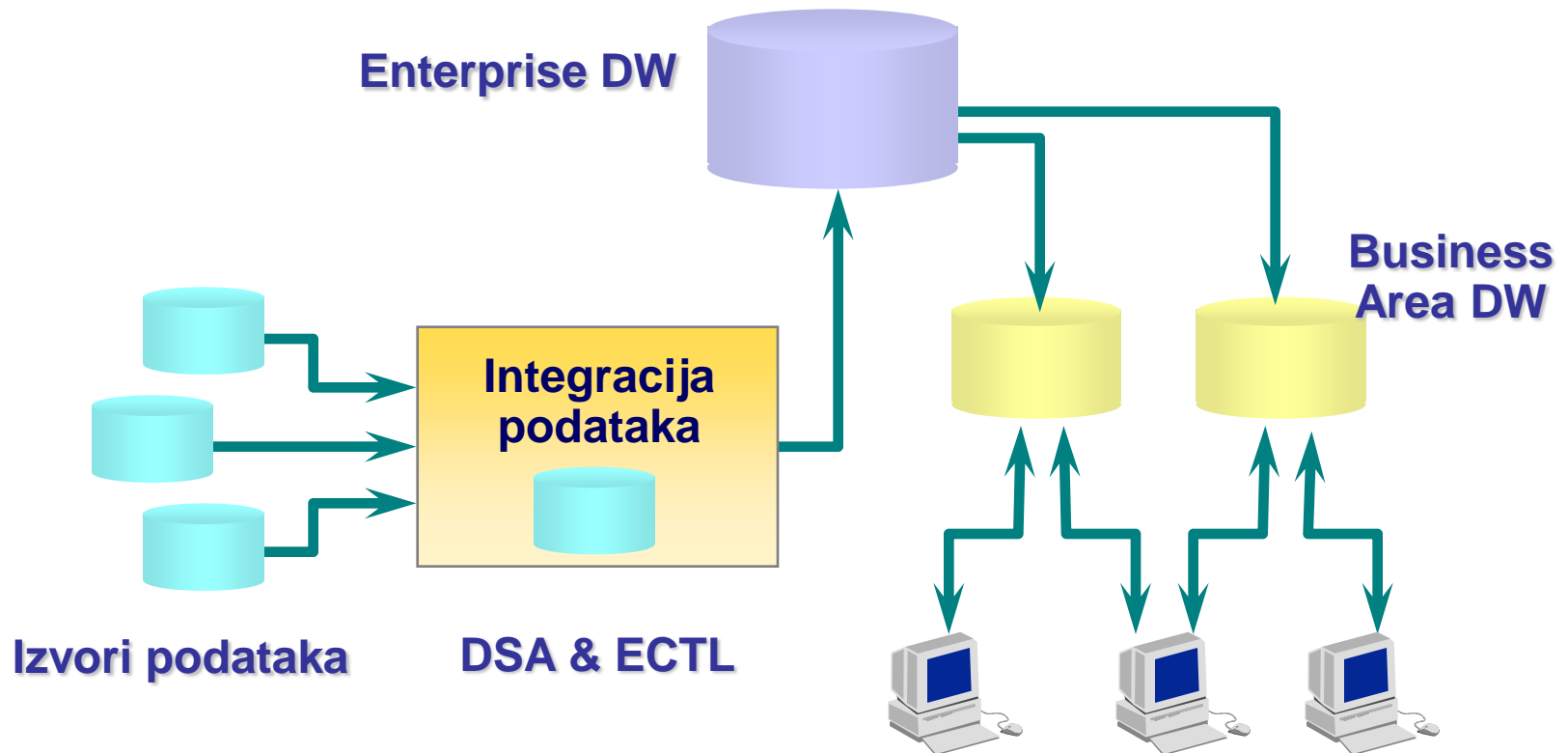
# Arhitektura DW sistema

---

- **Kreiranje i osvežavanje DW**
  - ETL (ECTL) proces
  - softverska podrška za zadatke
    - **Extraction (ekstrakcija)**
      - selektovanje (izdvajanje) podataka iz različitih izvora
    - **(C)leaning & Transformation (transformacija)**
      - validacija, pročišćavanje, integracija i vremensko označavanje podataka
    - **Loading (punjenje)**
      - punjenje DW baze podataka

# Arhitektura DW sistema

- Primer
  - moguće različite kombinacije u arhitekturi DW



# Arhitektura DW sistema

---

- **Klasifikacija DW sistema**
  - prema opsegu pokrivenosti tema
    - **Enterprise Data Warehouse (EDW)**
      - pokriva celokupno poslovanje
    - **Data Mart (DM)**
      - pokriva samo jedan segment poslovanja
        - » uobičajeno, jednu temu





# Arhitektura DW sistema

- Enterprise DW / Data Mart**

Osobina	Data Warehouse	Data Mart
Oblast	Poslovni sistem	Sektor (org. celina)
Teme	Više	Jedna
Izvori	Više	Manji broj (nekoliko)
Tipična veličina	100 GB - > N TB	< 100 GB
Vreme implementacije	Meseci - godine	Meseci

# Arhitektura DW sistema

---

- **Enterprise Data Warehouse (EDW)**
  - **Korporativni DW**
  - pokriva celokupno poslovanje
    - obuhvata podatke vezane za sve teme poslovanja
    - razvija se, uobičajeno, inkrementalnim pristupom
    - predstavlja jedan izvor podataka za celokupan menadžment poslovnog sistema
    - "sinhronizuje" podatke iz svih izvora podataka poslovnog sistema
    - može biti osnova za izgradnju pojedinačnih Data Mart sistema

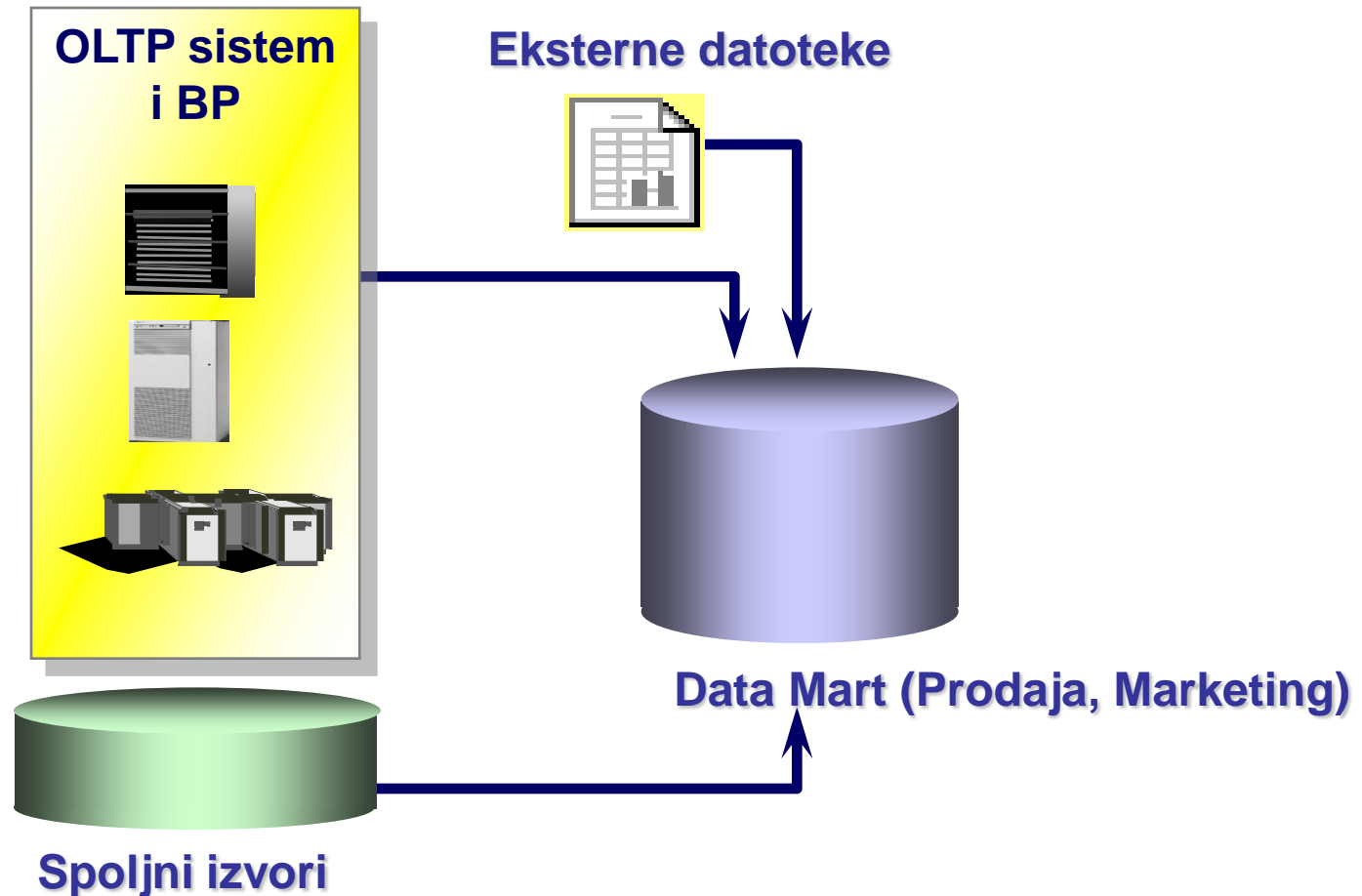
# Arhitektura DW sistema

---

- **Data Mart (DM)**
  - **Tematski DW**
  - pokriva samo jedan segment poslovanja
    - obuhvata podatke vezane za jednu temu poslovanja
    - može predstavljati pilot projekat za realizaciju obimnijeg DW sistema
    - može biti realizovan kao
      - nezavisni Data Mart
        - » izolovan od drugih DW sistema
      - zavisni Data Mart
        - » naslonjen na druge DW sisteme

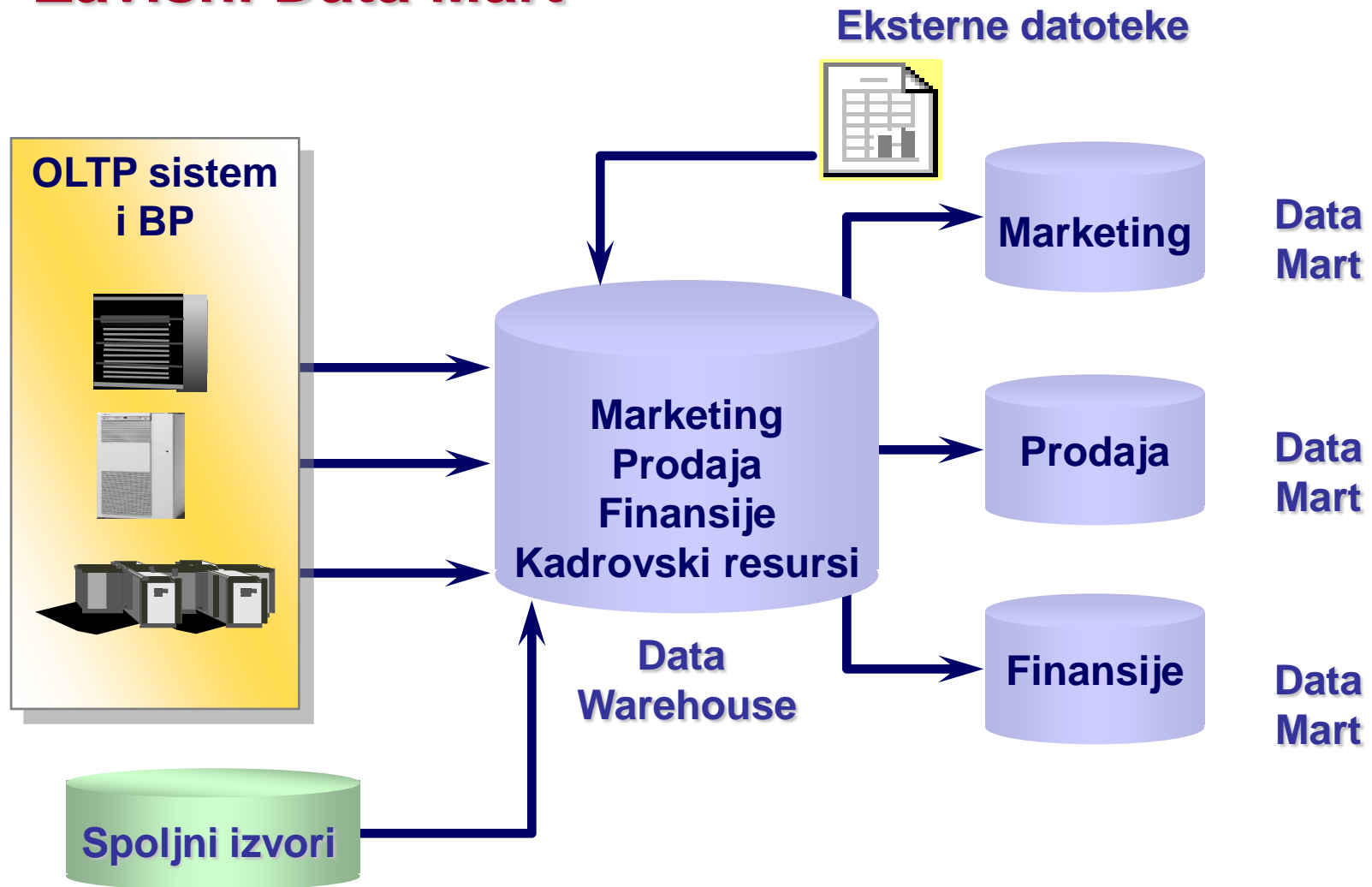
# Arhitektura DW sistema

- **Nezavisni Data Mart**



# Arhitektura DW sistema

- Zavisni Data Mart**



# Arhitektura DW sistema

---

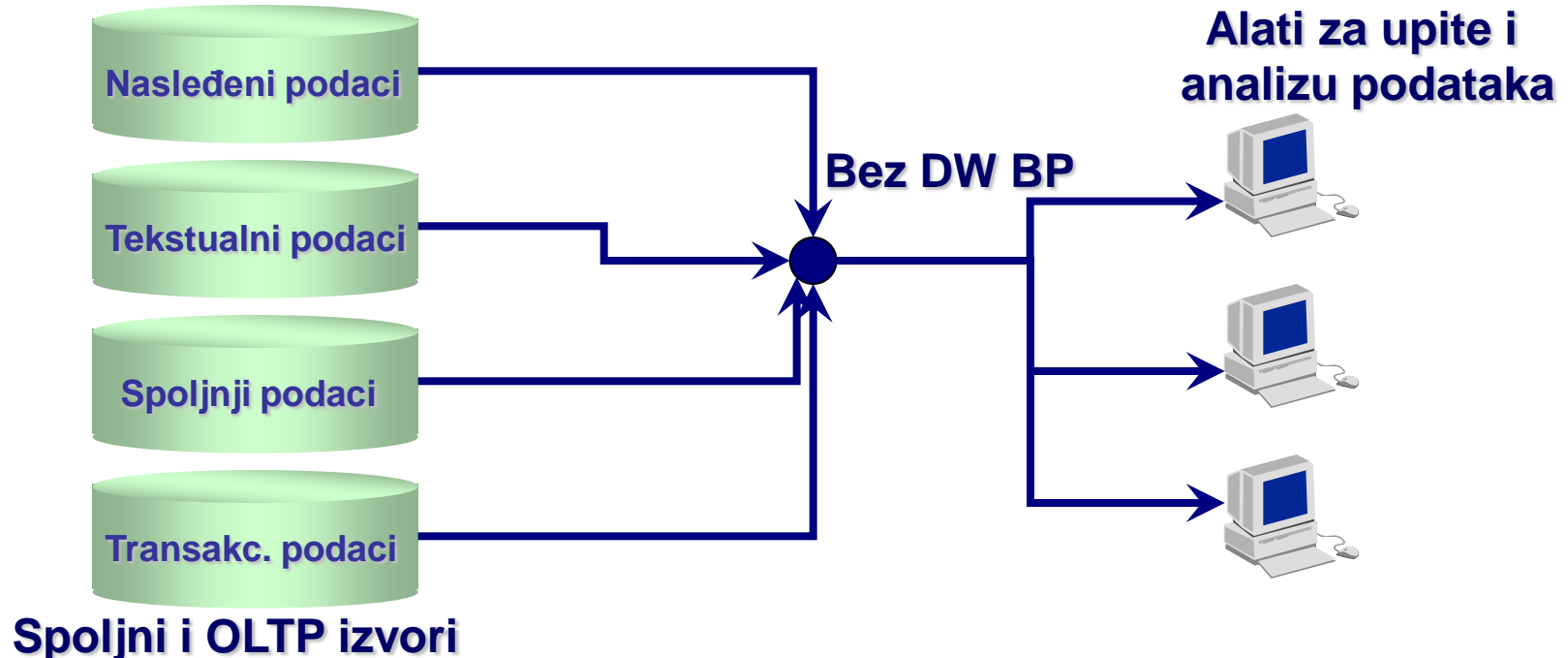
- **Vrste arhitektura DW sistema**
  - virtuelna DW arhitektura
  - upakovana DM arhitektura
  - arhitektura izvedenog DM
  - arhitektura povezanog DM
  - arhitektura sa centralnim DW

# Arhitektura DW sistema

- **Virtuelna DW arhitektura**

- **"Virtual DW" Architecture**

- ne postoji posebna DW BP
    - obezbeđen direktan pristup OLTP sistemu, korišćenjem (jednostavnijih) klijentskih alata za upite i analizu podataka



# Arhitektura DW sistema

---

- **Virtuelna DW arhitektura**
  - **"Virtual DW" Architecture**
    - prednosti
      - mala dodatna ulaganja u IT - hardver, softver i ljudski resursi
      - nema posebnog ECTL softvera
      - jednostavna upotreba klijentskih alata za upite i analizu podataka
      - mala ulaganja u obuku krajnjih korisnika
    - nedostaci
      - ne postoje istorijski podaci
      - ne postoje agregirani podaci
      - ne postoje centralizovani meta podaci sa opisom semantike
      - ne postoje procedure za pročišćavanje i integraciju podataka
    - privremeno i kratkotrajno rešenje



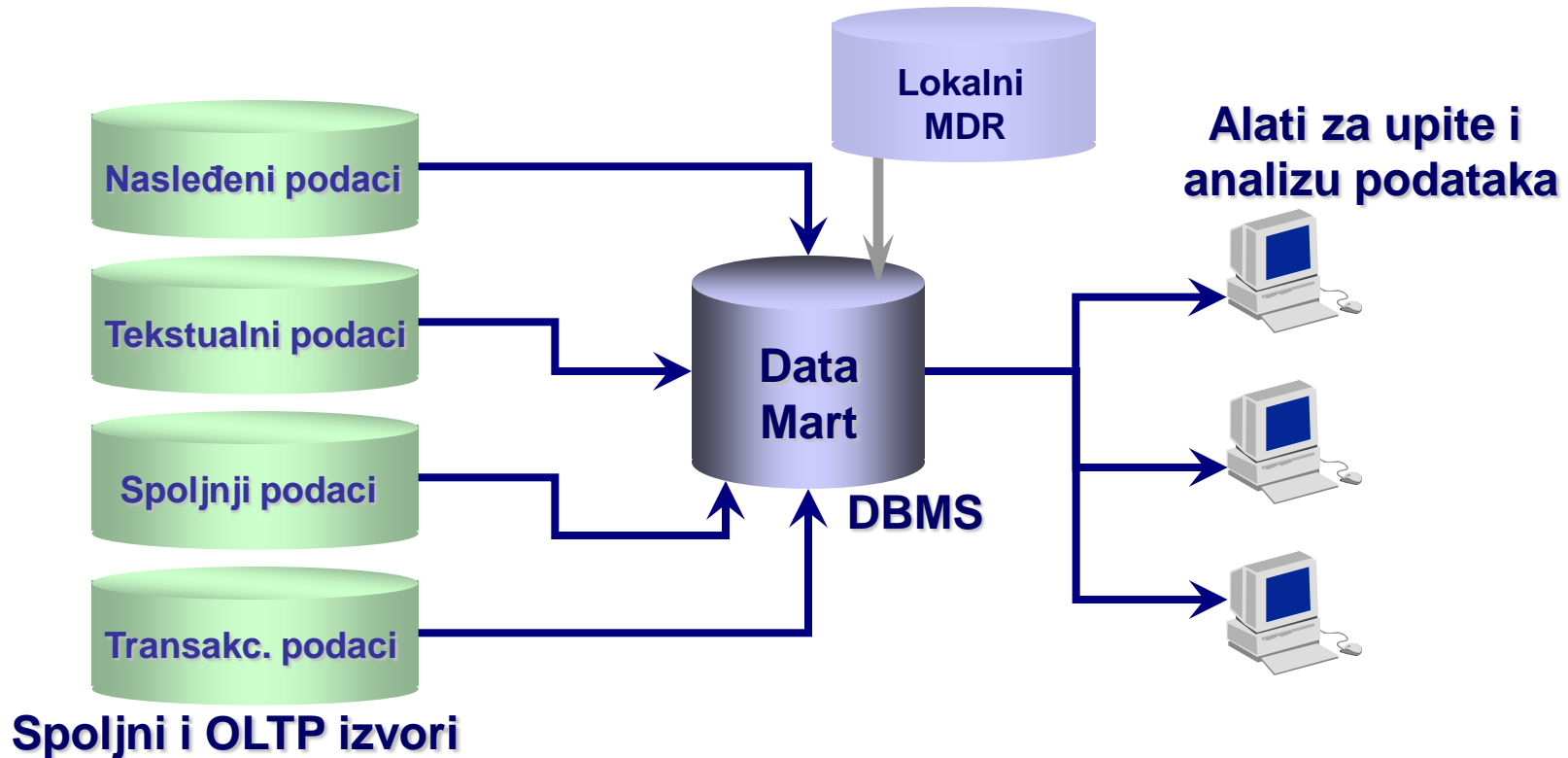
# Arhitektura DW sistema

---

- **Upakovana DW arhitektura**
  - **"Data Mart in a Box" Architecture**
  - "upakovani" softverski proizvod koji obezbeđuje
    - izgradnju DW BP koristeći različite izvore podataka
    - pristup DW BP upotrebom jednostavnih alata za upite i analizu podataka ("user friendly" tools)
    - izgradnju lokalnog MDR
      - sa definicijama podataka iskazanim terminologijom aplikativnog domena

# Arhitektura DW sistema

- **Upakovana DW arhitektura**
  - "Data Mart in a Box" Architecture



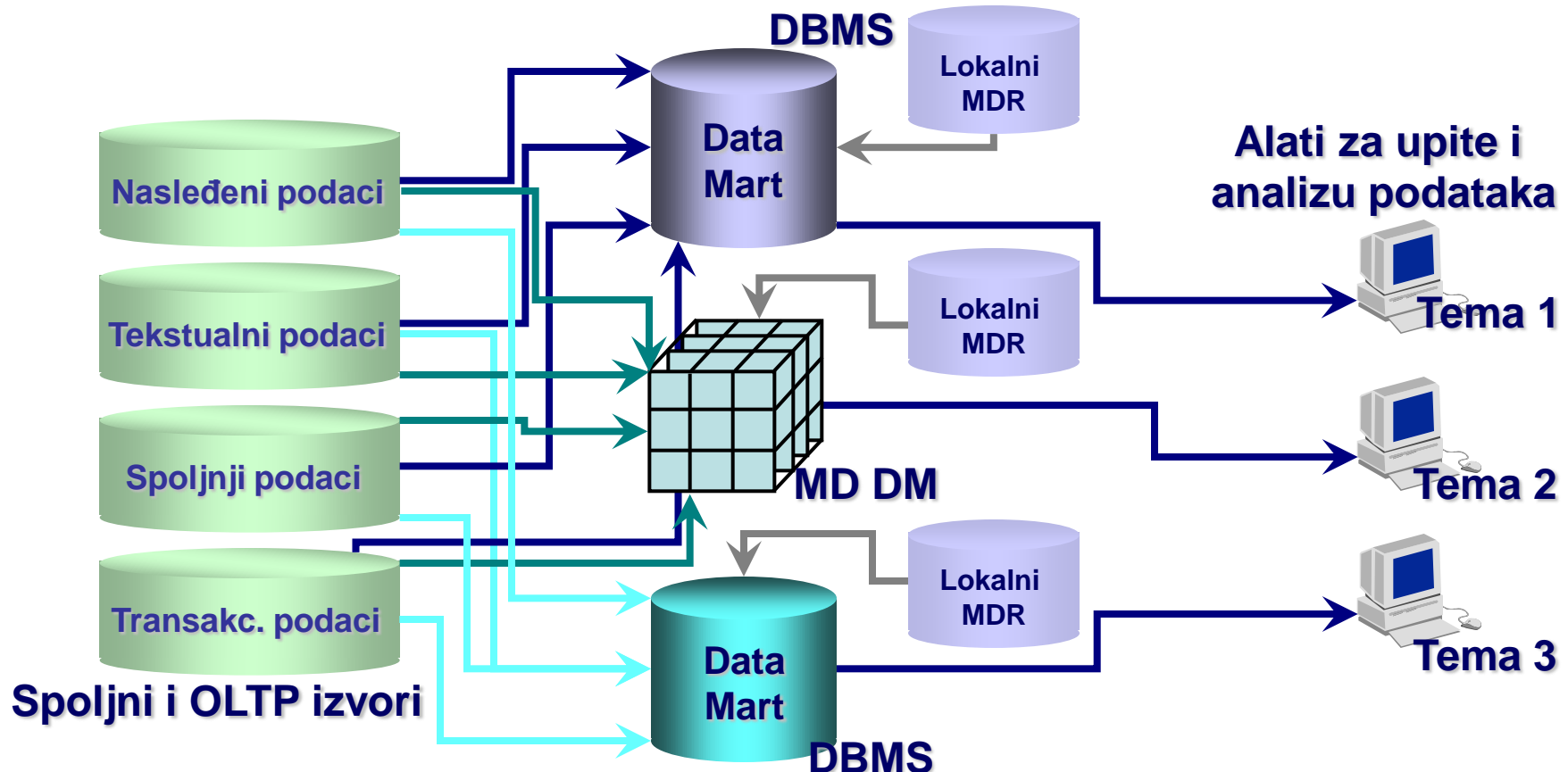
# Arhitektura DW sistema

---

- **Upakovana DW arhitektura**
  - **"Data Mart in a Box" Architecture**
    - prednosti
      - eliminiše potrebu direktnog komuniciranja sa OLTP BP
    - nedostaci
      - favorizacija nekontrolisane "eksplozije" nezavisnih, neintegrisanih DM, često
        - » s nepotrebnom redundancijom podataka
        - » nabavljenih od strane različitih proizvođača
      - nepostojanje zajedničke osnove meta podataka na nivou organizacije i različitih tema poslovanja
        - » za iskazivanje semantike podataka i pravila poslovanja
      - favorizacija preuzimanja prljavih podataka iz različitih izvora

# Arhitektura DW sistema

- **Upakovana DW arhitektura**
  - **"Data Mart in a Box" Architecture**
    - eksplozija nezavisnih DM



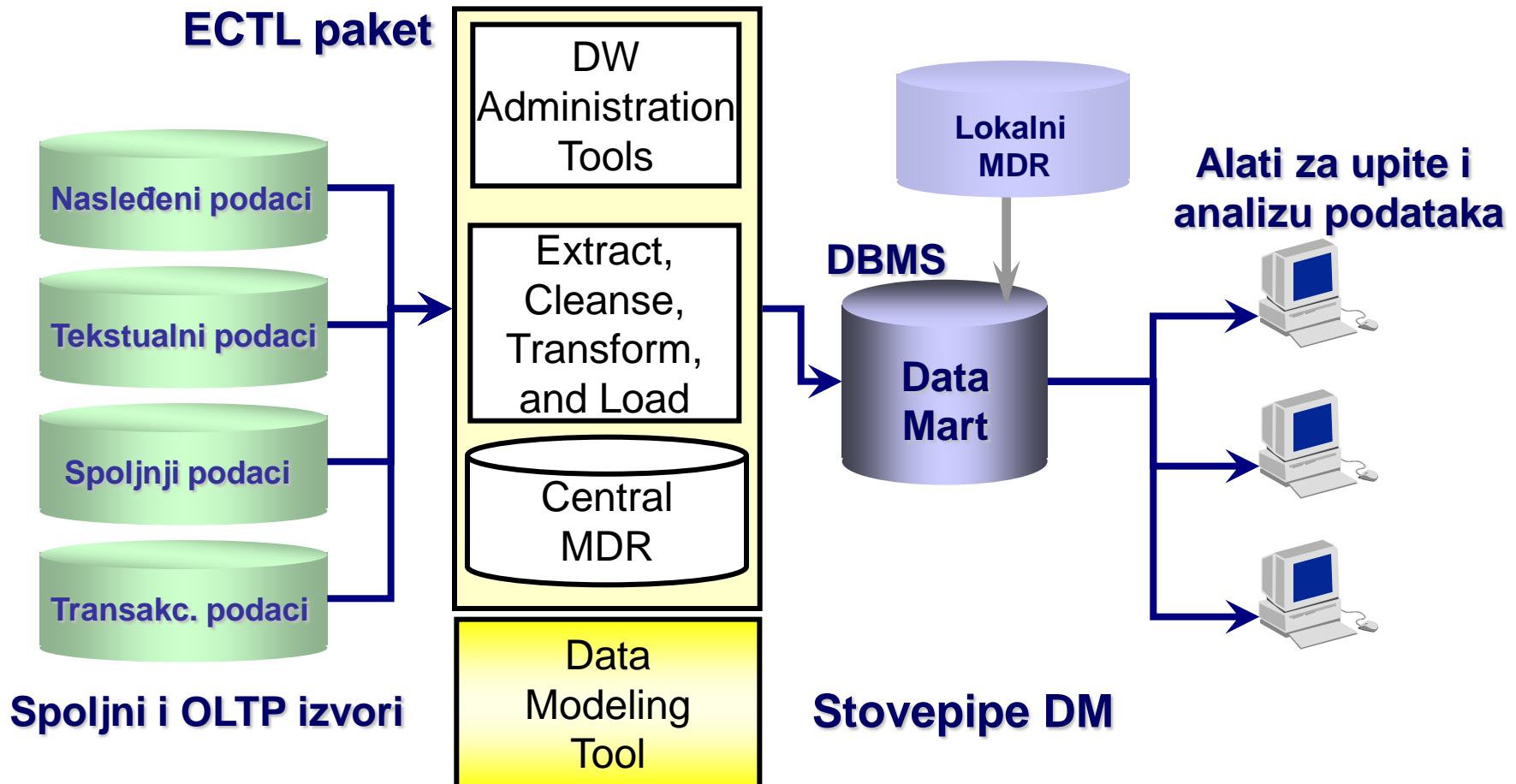
# Arhitektura DW sistema

---

- **Arhitektura izvedenog DM**
  - **"Stovepipe Data Mart" Architecture**
  - upakovana DM arhitektura, opremljena jednim ECTL softverskim paketom i DSA međupodručjem sa funkcijama
    - ekstrakcija, pročišćavanje, transformacija i punjenje podataka
    - agregiranje i sumiranje podataka
    - kreiranje i održavanje centralizovanog MDR (CMDR)
    - obezbeđenje administratorskih funkcija DW sistema
    - obezbeđenje interfejsa prema alatu za modelovanje DW BP

# Arhitektura DW sistema

- **Arhitektura izvedenog DM**
  - "Stovepipe Data Mart" Architecture



# Arhitektura DW sistema

---

- **Arhitektura izvedenog DM**
  - **"Stovepipe Data Mart" Architecture**
    - prednosti
      - upotreba jedinstvenog ECTL softvera
      - uvođenje centralizovanog MDR
    - nedostaci
      - nikakva ili slaba integrisanost CMDR i lokalnih MDR
        - » proizvođači "stovepipe DM" softvera često ne pružaju mehanizme za uspostavljanje veze između CMDR i lokalnih MDR
      - favorizacija izgradnje međusobno nezavisnih i neintegrisanih DM
      - podrška organizacionih potreba na nivou jedne teme, ili jednog tipa radnog mesta
        - » nedostatak podrške potreba celokupne organizacije

# Arhitektura DW sistema

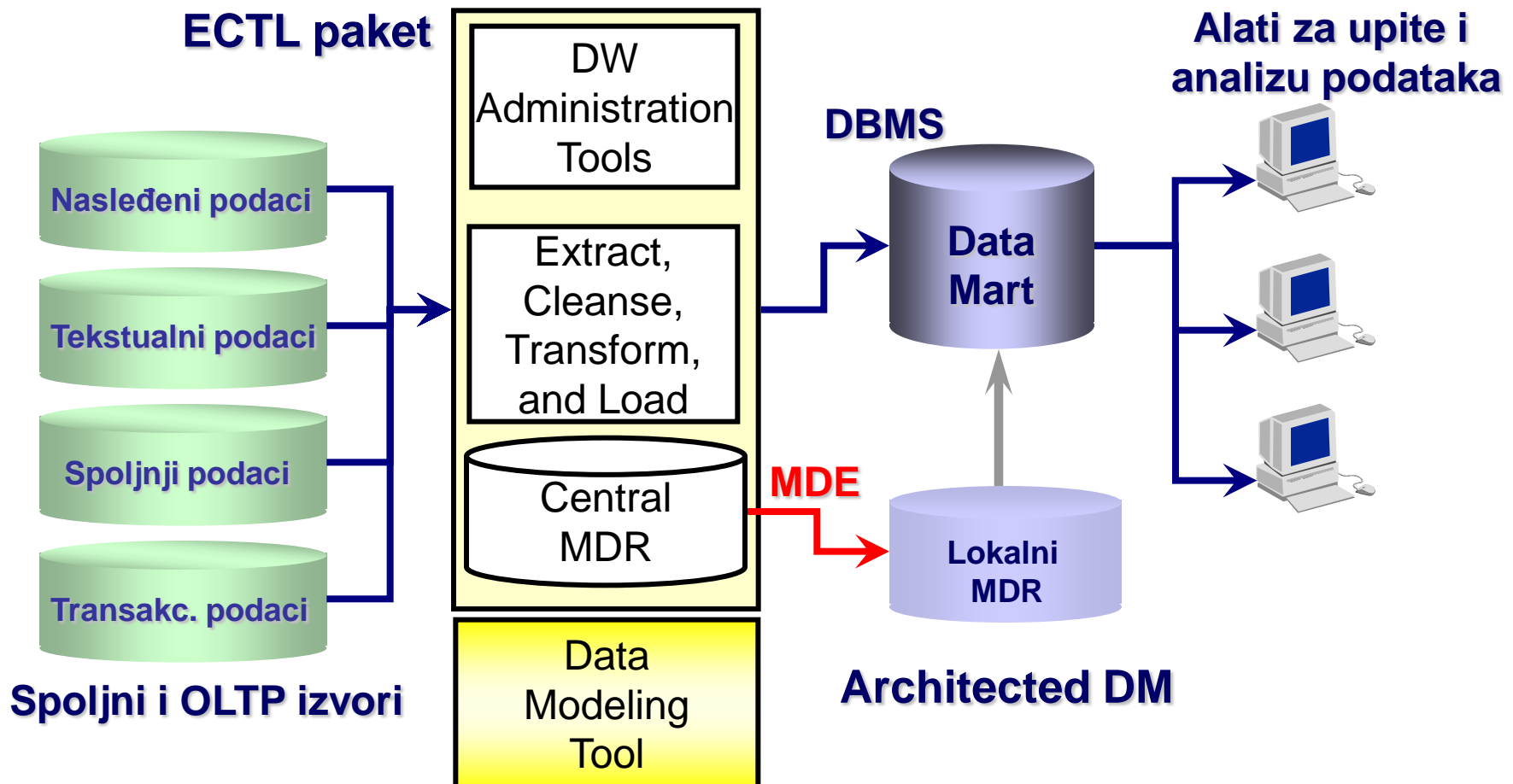
---

- **Arhitektura povezanog DM**
  - **"Architected Data Mart" Architecture**
  - arhitektura izvedenog DM, opremljena softverom za razmenu meta podataka
    - **Meta Data Exchange (MDE) Component**
    - obezbeđenje usaglašenosti CMDR i lokalnih MDR
    - CMDR postaje centralno mesto znanja, ugrađenog u DW sistem



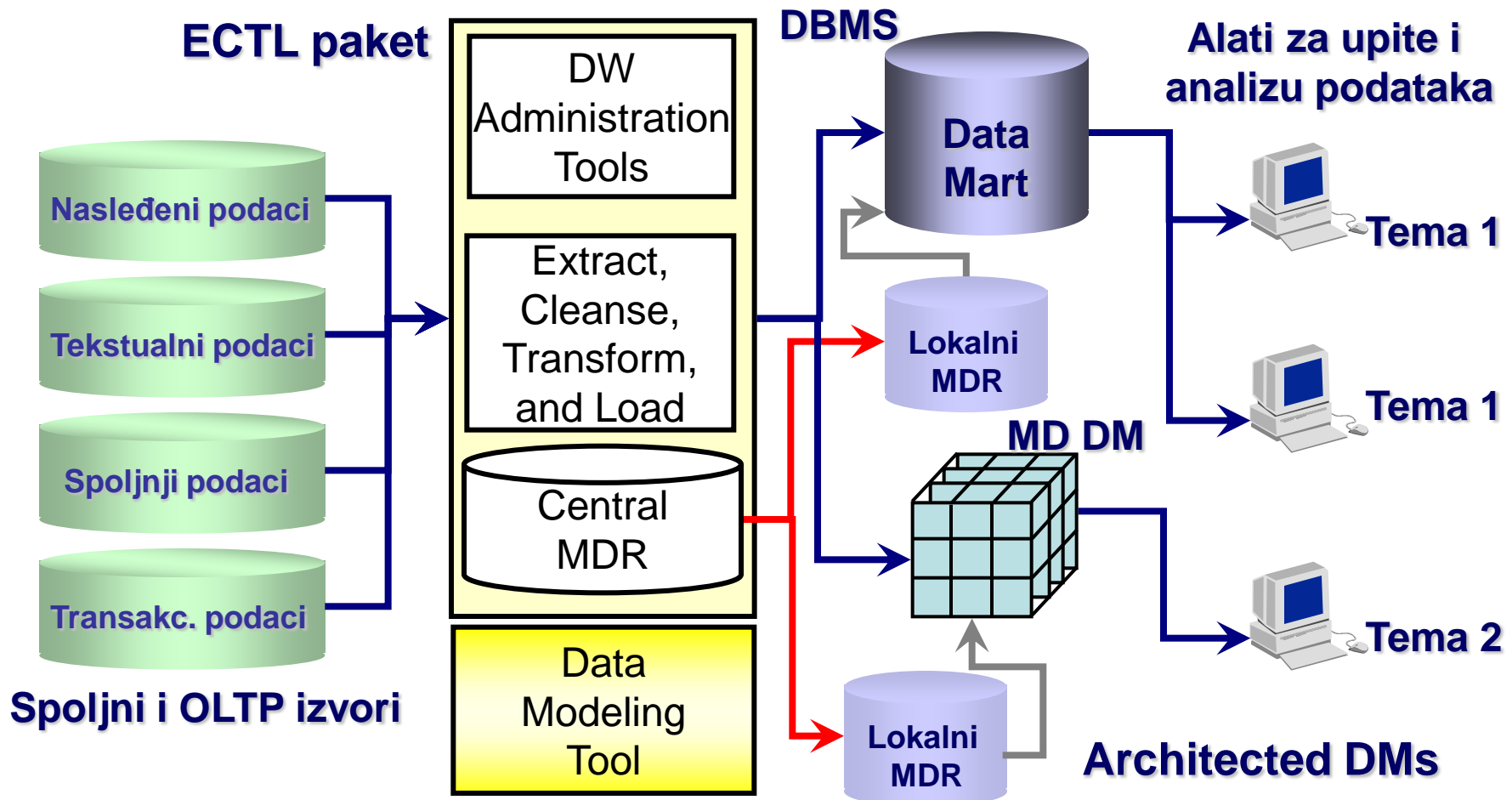
# Arhitektura DW sistema

- **Arhitektura povezanog DM**
  - "Architected Data Mart" Architecture



# Arhitektura DW sistema

- **Arhitektura povezanog DM**
  - "Architected Data Mart" Architecture



# Arhitektura DW sistema

---

- **Central Meta Data Repository (CMDR)**
  - "single version of the truth"
  - sadrži
    - definicije izvornih podataka na nivou organizacije
    - opise semantike podataka
    - specifikacije logičkih i fizičkih modela ciljne BP
    - opise izvora podataka
    - pravila preslikavanja izvornih u ciljne podatke
    - pravila pročišćavanja podataka
    - pravila transformacije podataka
    - opise procedura za agregiranje i sumiranje podataka
  - nažalost, ne postoji opšte prihvaćeni standard za opis MDR modela

# Arhitektura DW sistema

---

- **Arhitektura sa centralnim DW**
  - **"Central Data Warehouse" (CDW) Architecture**
  - jedinstvena i integrisana BP DW sistema
  - čuva analitičke podatke
    - može do nivoa elementarnih ("dnevni") transakcija
  - predstavlja izvor usaglašenih podataka nivoa cele organizacije
  - analitičari podataka je koriste za upite i izveštaje nad analitičkim (atomičnim) podacima
  - odvojena od OLTP BP
    - ne koristi se za direktno ažuriranje podataka putem elementarnih transakcija
  - osnova za formiranje i osvežavanje BP povezanih DM

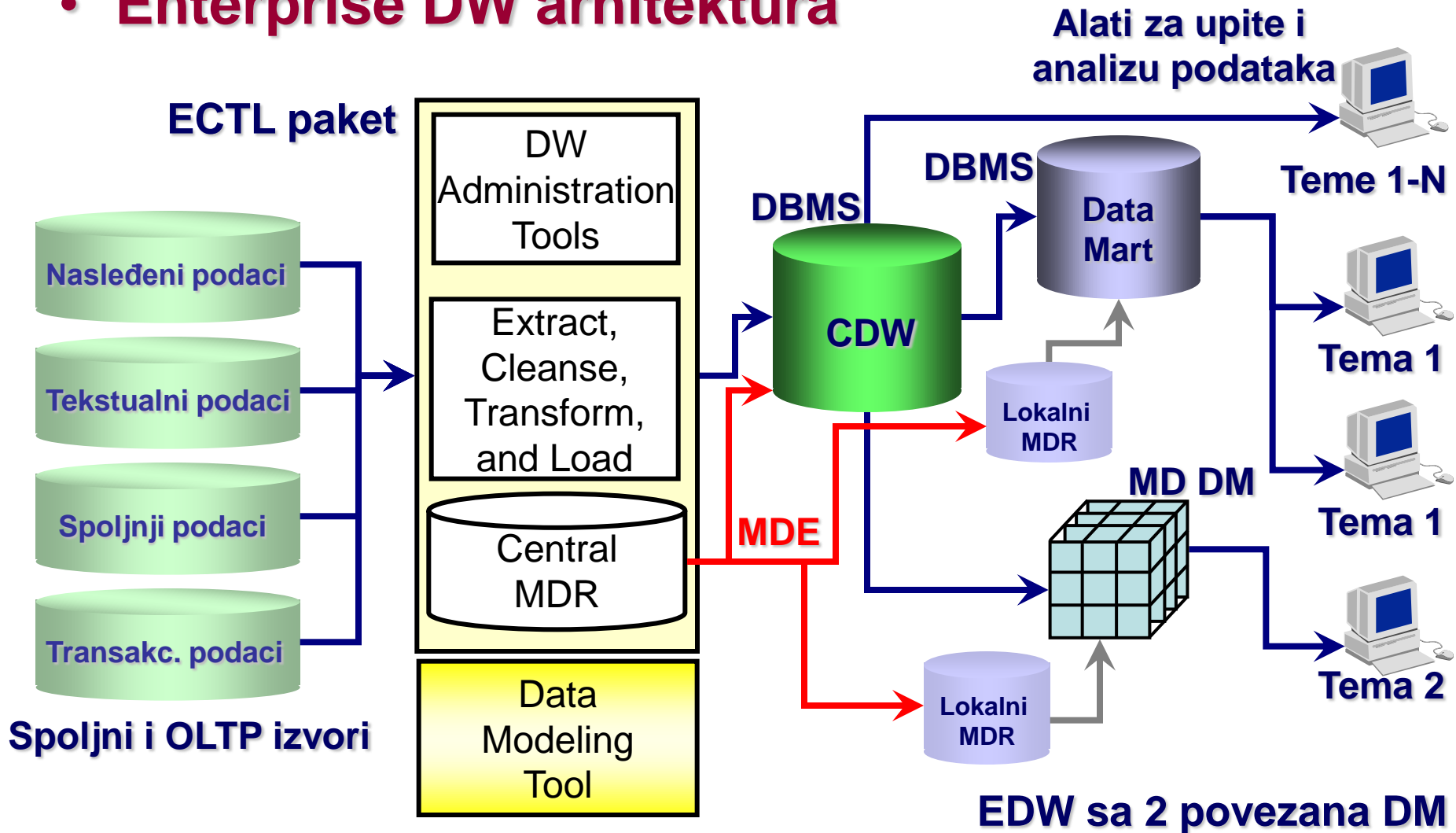
# Arhitektura DW sistema

---

- **Arhitektura korporativnih DW**
  - **"Enterprise DW (EDW)" Architecture**
  - karakteristike
    - više izvora podataka
    - "Off-the-Shelf" ECTL softverski paket
    - Central Meta Data Repository (CMDR)
    - Meta Data Exchange (MDE) Component
    - Central Data Warehouse (CDW)
    - primena više povezanih DM
    - centralno upravljanje DW sistemom, preko CMDR
    - primena različitih alata za upite i analizu podataka
    - Web orijentisana aplikativna podrška

# Arhitektura DW sistema

## • Enterprise DW arhitektura



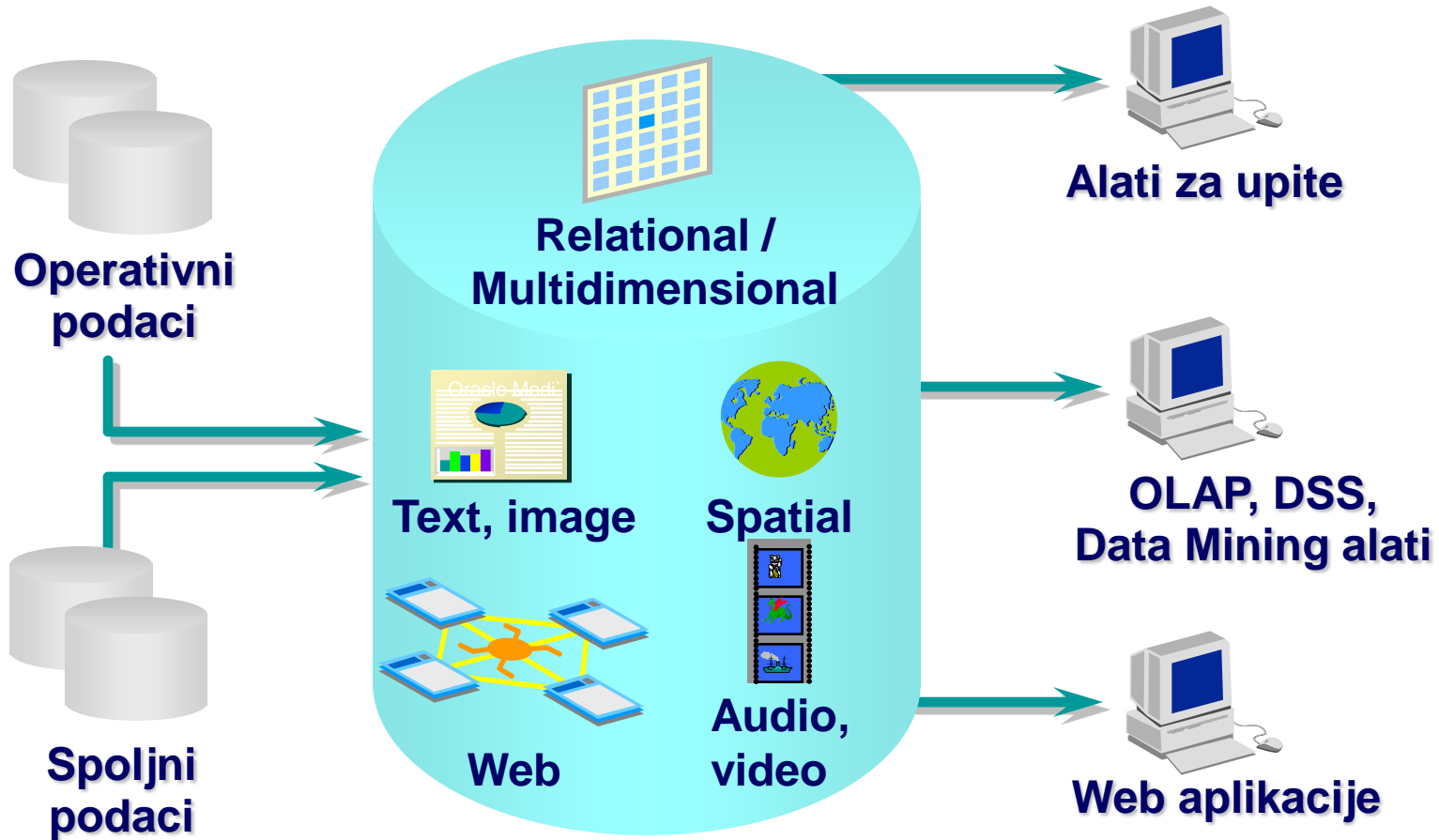
# Sadržaj

---

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

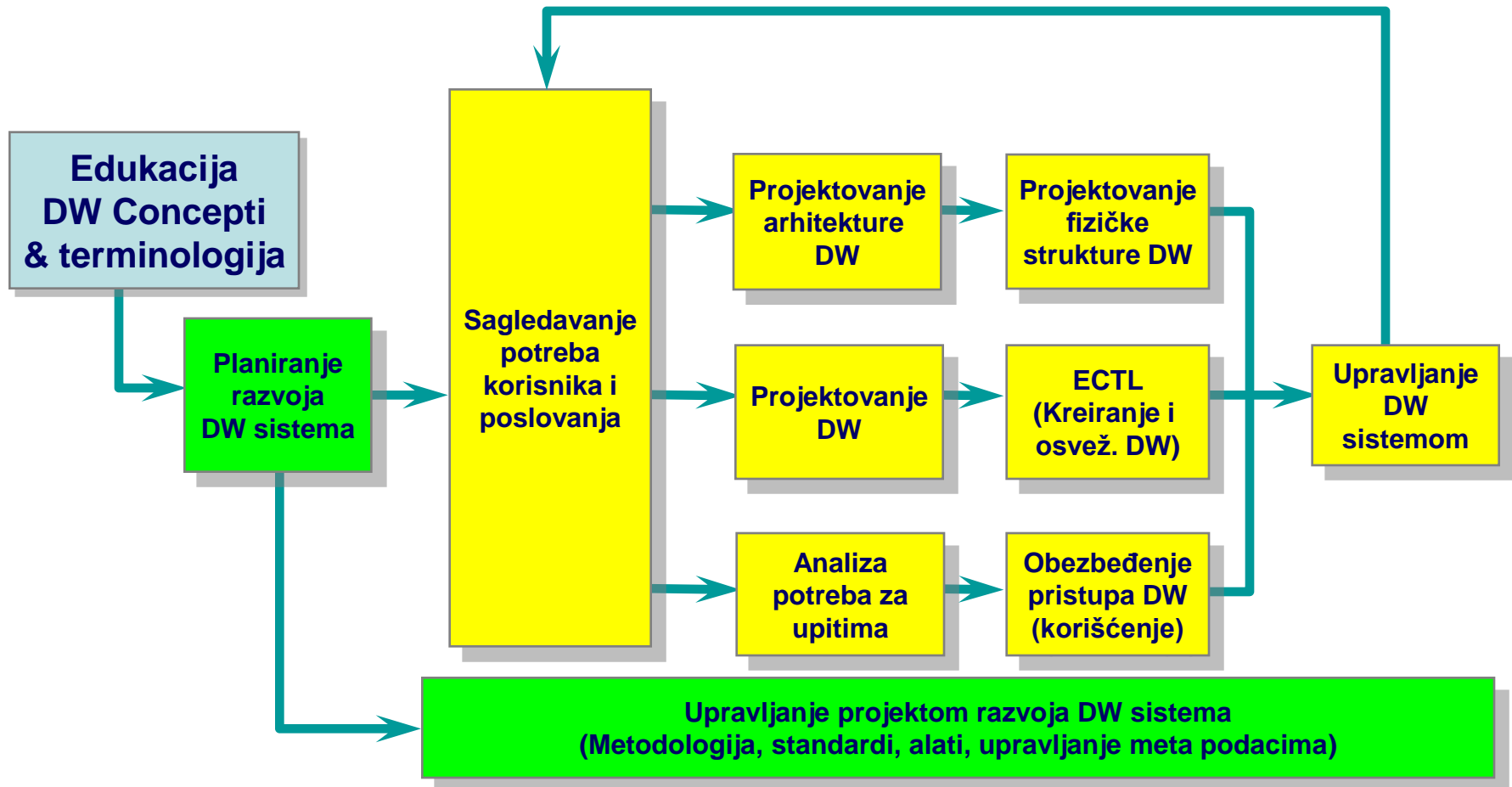
# Razvoj DW sistema

**Različiti izvori**    **Podaci različitog tipa**    **Alati različite namene**





# Razvoj DW sistema

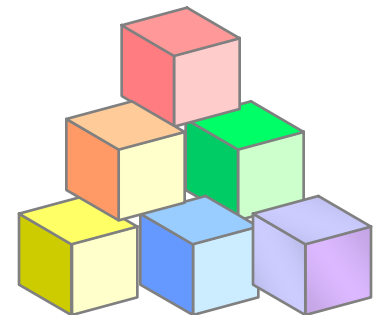


# Razvoj DW sistema

- **Projektovanje DW**

- **opšti zahtevi**

- **primena odgovarajuće metodologije**
      - ciljevi i postupci (procesi) projektovanja
      - tehnike projektovanja
      - projektantski softverski alati
      - projektantski standardi (uputstva)
    - **primena opštih principa softverskog inženjerstva**
    - **uvažavanje namene i prirode DW sistema**
      - različite od namene i prirode OLTP sistema

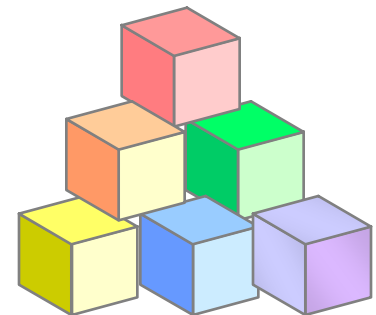


# Razvoj DW sistema

- **Projektovanje DW**

- **poslovni zahtevi**

- struktura i sadržaj DW se, u opštem slučaju, razlikuje od OLTP BP
      - zahtevi za obezbeđenje analiza podataka
      - orijentacija podataka u DW na teme
        - » identifikacija tema poslovanja
        - » definisanje veza između tema poslovanja
        - » identifikacija i imenovanje atributa tema

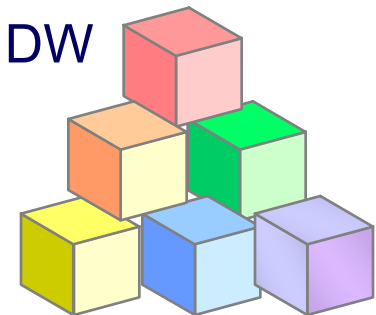


# Razvoj DW sistema

- **Projektovanje DW**

- **projektantski zahtevi**

- projektovanje Enterprise DW je, praktično, **iterativan proces**
    - metodologija treba da bude zasnovana na
      - iterativnom (evolutivnom) pristupu i
      - inkrementalnom pristupu
        - » **sigurna i proverena**
        - » **upravljiva**
        - » **preporučljiva**
    - postoje odgovarajući CASE alati koji mogu da podrže iterativno i inkrementalno projektovanje DW



# Razvoj DW sistema

---

- **Projektovanje arhitekture DW**

- zahtev

- analiza i uvažavanje strateških ciljeva i potreba organizacije
- okvir za zadovoljenje specificiranih korisničkih zahteva
- fleksibilnost i proširivost
- dugovečnost
- opravdanost investicije u DW sistem

- problemi

- puno projekata DW sistema propada usled zahteva da razvoj DW sistema bude brz i jeftin
- posledica: izbor neodgovarajuće arhitekture s obzirom na poslovne zahteve

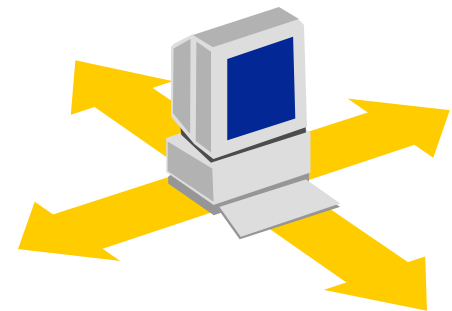
# Wrong DW Architecture

---

- **Projektovanje arhitekture DW**
  - arhitekture koje, na duži period vremena, mogu pokazati loše osobine
    - “Virtual” Data Warehouse
    - “Data Mart in a Box”
    - “Stovepipe” Data Marts

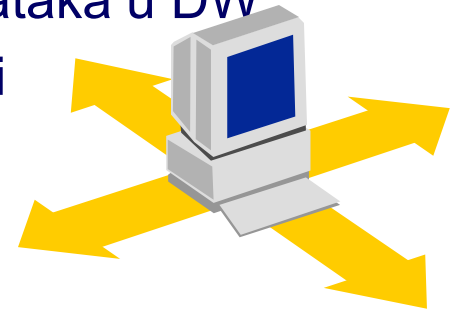
# Razvoj DW sistema

- **Projektovanje fizičke strukture DW**
  - zahtev
    - efikasna podrška velikih BP (>TB)
    - arhitektura servera DW BP
      - karakteristike hardvera i OS
    - mehanizmi SUBP
      - fizičko particioniranje BP
      - paralelno izvršavanje upita i optimizacija upita
      - tehnike indeksiranja
        - » bitmap, join, project i bit-slice indeksne strukture
      - tehnike replikacije BP
        - » materijalizovani pogledi



# Razvoj DW sistema

- **Kreiranje i osvežavanje DW - ECTL**
  - zahtev
    - odgovarajući algoritmi, tehnike i alati za ECTL zadatke
    - odgovarajući mehanizmi SUBP
      - tehnike predaje podataka (export)
      - tehnike preuzimanja podataka (import)
      - tehnike transfera podataka (na fizičkom i logičkom nivou)
      - tehnike masovnog ažuriranja podataka
      - tehnike optimizacije masovnih upita i izračunavanja nad podacima
    - obezbeđenje odgovarajućeg kvaliteta podataka u DW
    - obezbeđenje zadovoljavajućih performansi





# Razvoj DW sistema

---

- **Kreiranje i osvežavanje DW - ECTL**

- problemi

- problem "priljavih podataka" (the "dirty data" problem)
  - podaci u operativnim - nasleđenim (OLTP) BP često sadrže veliki broj nedostajućih, oštećenih ili nekonzistentnih vrednosti
- primeri
  - više vrednosti "upakovanih" kao jedna vrednost atributa
  - jedna logička vrednost atributa podeljena na više vrednosti različitih atributa
  - ista vrednost atributa, zapisana na više različitih načina
    - » "spelling errors" (primer u srpskom: "dj" / "đ")
    - » upotreba različitih kodnih rasporeda ili pisama (ćirilica / latinica)
  - nekonzistentni nazivi ili čak adrese konkretnih entiteta
  - nekonzistentne vrednosti ključeva istih entiteta
- statistički, čak do 20% vrednosti u BP sadrži prijave podatke

# Razvoj DW sistema

---

- **Kreiranje i osvežavanje DW - ECTL**
  - zadaci
    - eliminacija ili bar ublažavanje problema "prijavih podataka"
    - centralna tačka za upravljanje pristupom svim izvorima podataka
    - generisanje i održavanje centralnog repozitorijuma meta podataka

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - zahtev
    - primena odgovarajućih alata za korišćenje DW BP
  - zadaci – obezbeđenje
    - prezentacije podataka
    - opšte podrške pristupa DW
    - zadavanja upita i generisanja izveštaja
    - analize podataka
    - istraživanja podataka
    - podrške procesa odlučivanja

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - **prezentacija podataka**
    - GUI, Windows, Web i štampane forme
    - u alfanumeričkom, grafičkom, ili grafikonskom obliku
  - **opšta podrška pristupa DW**
    - aplikacije za rad sa DW bazom podataka
      - Windows orijentisane
      - Web orijentisane, bazirane na koncepciji Portala
      - dinamički konfigurabilne
        - » prema potrebama, pravima pristupa i željama korisnika
        - » mogućnost lokalizacije i parametrizacije softvera

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - **zadavanje upita i generisanje izveštaja**
    - inteligentni i GUI orijentisani editori upita
    - inteligentni i GUI orijentisani editori formi za izveštaje
    - podrška izgradnje različitih tipova izveštaja
      - "jednostavnih" i "master-detail" ("zaglavlje-stavke")
      - "drill-down" ("dubinski orijentisanih")
      - egzaktnih
      - prognostički orijentisanih

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - analiza podataka
    - upotreba On-line Analytical Processing (OLAP) alata
      - upita i izračunavanja nad višedimenzionalnim strukturama
        - » strukture tipa hiperkocke
    - upotreba matematičkih modela analize podataka
      - linearno programiranje
      - regresione analize
      - analize vremenskih serija

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - istraživanje podataka
    - upotreba Data Mining alata
      - rezonovanje (izvođenje zaključaka) na osnovu velike količine podataka (iz DW)
      - podrška formalnih metoda rezonovanja
        - » veštačka inteligencija  
(otkrivanje znanja i mašinsko učenje)
      - podrška statističkih metoda
        - » istraživačke analize podataka
        - » soft computing

# Razvoj DW sistema

---

- **Obezbeđenje pristupa DW**
  - **podrška procesa odlučivanja**
    - upotreba inteligentnih DSS (ekspertskih sistema) za podršku
      - identifikacije problema odlučivanja
      - identifikacije vrednih alternativa pri odlučivanju
      - primene matematičkih modela odlučivanja
      - ocenjivanje alternativa pri odlučivanju
      - izbor alternative pri odlučivanju
      - simulaciju primene izabrane alternative
      - praćenja i analize efekata primene izabrane alternative



# Sadržaj

---

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Trendovi u razvoju DW sistema

---

- **On-Line Enterprise Reporting (OLER)**
  - noviji pogled na primenu DW sistema u poslovanju
  - motivacija
    - neophodnost sistematične podrške izveštajne funkcije kompletne organizacije
      - ne samo strateškog i taktičkog menadžmenta organizacije
    - koristeći principe izgradnje DW sistema
  - DW sistem u funkciji repozitorijuma podataka kompletnog organizacionog sistema
    - obezbeđenje visokog nivoa integrisanosti podataka organizacije
      - često skoro operativnih ili baš operativnih podataka
      - u situacijama kada je to nemoguće obezbediti na nivou primenjenih OLTP sistema

# Trendovi u razvoju DW sistema

---

- Pomak u pogledu na definiciju DW sistema
  - DW (1993), Bill Inmon
    - subject oriented
    - integrated (data copies)
    - time variant (time series)
    - non volatile
      - primarno u upotrebi u DSS
  - DW (200x), trendovi
    - near real-time
    - event oriented
    - business process oriented
    - central source data
    - za upotrebu u celokupnom unutrašnjem i spoljnjem poslovanju organizacije

# Trendovi u razvoju DW sistema

---

- **Near Real-Time DW**

- **DW u skoro realnom vremenu**

- stalni pritisak skraćanja vremenskih rokova realizacije procesa poslovanja

- zahtev

- visoka granularnost vremenske dimenzije

- da zadovolji potrebe praćenja promena u skoro realnom vremenu

- da bi se obezbedio uvid u trenutno stanje poslovanja

- granularnost nivoa nedelje ili meseca (pa i dana) klasičnih DW je najčešće nezadovoljavajuća

- moguće rešenje

- iste transakcije u realnom vremenu koriste se i za ažuriranje OLTP BP i za ažuriranje centralne DW BP

# Trendovi u razvoju DW sistema

---

- **Event Data DW**
  - **DW sa operativnim podacima**
  - motivacija
    - potreba da menadžeri u procesu odlučivanja koriste operativne, transakcione podatke
    - za razliku od klasičnih DW i DSS, koji se temelje na agregiranim podacima

# Trendovi u razvoju DW sistema

---

- **Business Process Oriented DW**
  - **DW orijentisan na procese poslovanja**
  - motivacija
    - potreba zadovoljenja zahteva korisnika saglasno logici procesa poslovanja (uključujući operativne procese)
      - potreba memorisanja operativnih transakcionih podataka u DW sistemu
    - za razliku od klasičnih DW i DSS, koji se temelje na tematski orijentisanim agregiranim podacima
      - organizovanim kao podaci dimenzija i činjenica

# Trendovi u razvoju DW sistema

---

- **Centralized Source Data DW**
  - **DW kao centralizovani izvor (rezpozitorijum) podataka**
  - zahtev
    - OLER DW treba da sadrži sve potrebne izvorne (transakcije) podatke
    - za potrebe izveštavanja (uključujući i ad hoc izveštaje) i praćenje poslovanja (auditing)
  - takav DW sistem predstavlja integrisanu BP celokupne organizacije

# Trendovi u razvoju DW sistema

---

- **DW for use inside and outside the company**
  - **DW za upotrebu u celokupnom unutrašnjem i spoljnjem poslovanju organizacije**
  - motivacija
    - DW se posmatra kao izvor podataka za sve korisnike unutar organizacije, kao i za poslovne partnere
    - za razliku od klasičnih DW, koji su prevashodno bili namenjeni menadžerima
  - pomak u pogledu na OLTP i DW
    - DW, kao centralizovani izvor podataka, isključivo se koristi za sve vrste izveštaja i upita
      - umesto OLTP sistema
    - OLTP sistem ostaje isključivo u funkciji obuhvata i ažuriranja transakcionih podataka



# Sadržaj

- Motivacija nastanka DW sistema
- Konceptcija DW sistema
- Tematske karakteristike DW sistema
- Poređenje karakteristika DW i OLTP
- Arhitektura DW sistema
- Razvoj DW sistema
- Trendovi u razvoju DW sistema

# Pitanja i komentari

---



Kraj prezentacije

# Osnovni koncepti Data Warehouse sistema

---

*Sistemi skladišta podataka*