

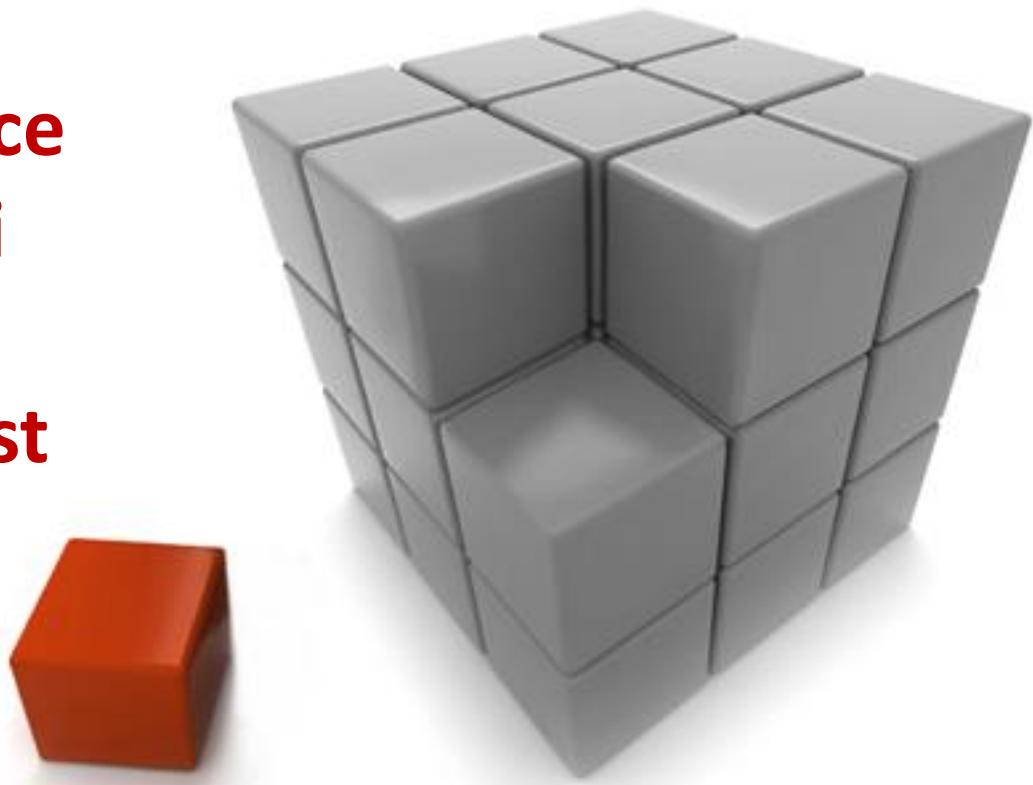


IN2 : **Business intelligence projekti u praksi i Data Science – pogled u budućnost**

Dušan Veličkovski, BI tim lider

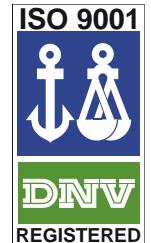
Jelena Lazarević, BI konsultant

Decembar, 2015



Agenda:

-  O IN2 Grupi
-  Šta je Business Intelligence i kome je namenjen
-  BI modeli i karakteristike BI projekata u praksi
-  Demonstracija BI4ERP rešenja na Tableau platformi
-  In2data – Data Science company
-  Primena SNA u praksi
-  Praktična primena Data Mininga u Osiguranju
-  Pitanja i diskusija



Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



O IN2 Grupi

- **Osnovana 1992. godine**
- **630 + zaposlenih**
- **U 5 najvećih SW kompanija u JI Evropi**
- **Prisutnost u regiji – gde su korisnici tu je IN2**
- **Kontinuirani rast**
- **Strategija – dugoročno partnerstvo**
- **Kompletna rešenja za branšu distribucije i prodaje-kada dođe vreme**
- **BI kao važan modul svake kompanije**

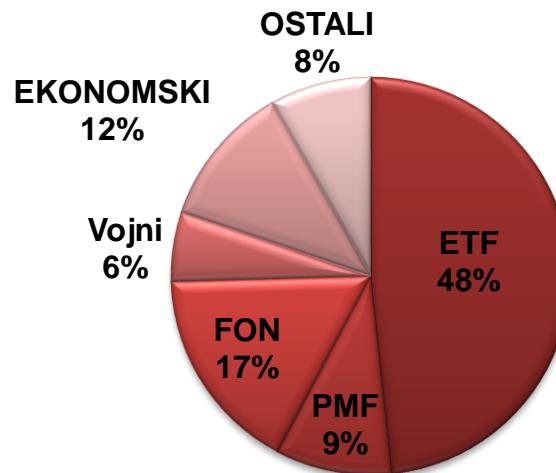
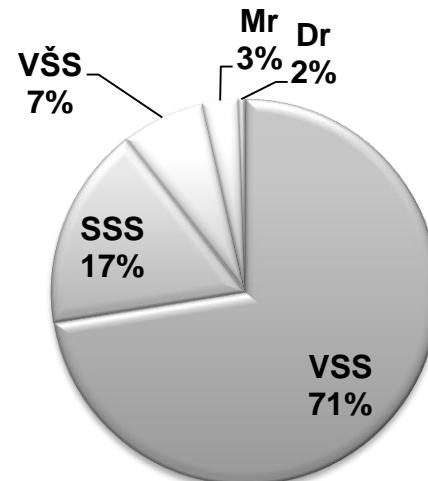


Profil IN2 grupe

Naši stručnjaci – ljudski kapital

- ❑ iskusni, visokoobrazovani i kreativni stručnjaci osnovna vrednost firme
- ❑ 630 zaposlenih u IN2 Grupi,
- ❑ preko 400 stručnih i poslovnih sertifikata
- ❑ konstantna edukacija zaposlenih u našem edukacionom centru
- ❑ prosečna starost 34 godine
- ❑ niska stopa odlazaka iz firme

Struktura zaposlenih po rangu obrazovanja



O IN2 Beograd

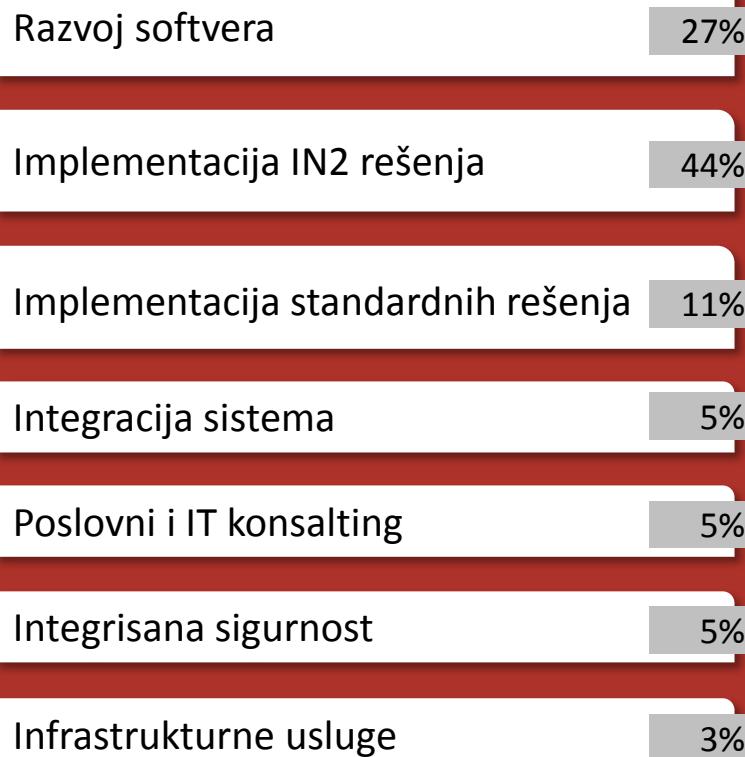
□ Osnovni podaci o IN2 Beograd:

- Trenutno 30+ ljudi, prihod oko 1M Eur
- IN2 - tržišni lider u Srbiji u sektoru dobrovoljnih penzijskih i investicionih fondova, kastodi banaka s aplikacionim sistemom **INvest2**
- Implementacija i podrška aplikacionog sistema **INsurance2 v3**
- Implementacija i podrška za Oracle **PSRM** (ETPM) rešenje
- Implementacija i podrška za WMS **Gold Stock** (SEYC)
- Jak Java tim, razvoj i održavanje **IN2 ADF** razvojnog okruženja
- Iskusan Oracle **Apex** tim
- **OBI i EPM** tim s rešenjima za osiguranje, prodaju i distribuciju
- **Data Science** tim

Poslovni model

PONUDA

Rešenja i usluge



TRŽIŠTA



Na čemu radimo trenutno ili pripremamo

□ IN2 u Konzumu

- Učesnik na projektu Oracle Retail, u budućnosti roll outs u BiH i Srbiji,
- Projekat KWS (Konzum Wholesale System) u Zagrebu
- Implementacija CRM i Campaign mgmt rešenja INsite2

□ Business Intelligence projekat u CROSO

- Business Intelligence sistem za Registraciju

□ Interesantna dolazeća budućnost

- Data Science i Big Data team u oblastima:
telekom, retail, finansije, javni sektor
- Digitalna transformacija poslovanja – dugoročne aktivnosti

IN2 reference u Srbiji

- ❑ **Narodna banka Srbije (SAGA)**
 - Implementacija Oracle E-Business Suite (ERP rešenje)

- ❑ **Centralni registar obaveznog socijalnog osiguranja (SAGA, SONECO)**
 - **Business Intelligence sistem za Registraciju**
 - Implementacija Oracle ETPM
 - Jedinstvena registracija na osiguranje
 - Praćenje uplate doprinosu od strane poslodavaca
 - Veza s PIO, RFZO, PU, MUP, APR, NSZ, Trezor

- ❑ **Atlantic grupa – implementacija WMS i OBI**

- ❑ **Naftna industrija Srbije - NIS BI Foundation (COMBIS) + EPM**

- ❑ **Komisija za zaštitu konkurentnosti (u toku - GIZ)**
 - Implementacija DMS rešenja (Alfresco DMS)

- ❑ **Telekom Srbija – (Nites, Fadata, Inplenion)**
 - Implementacija Oracle Hyperion EPM proizvoda



Agenda:

O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

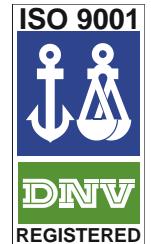
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



ORACLE® Platinum
Partner

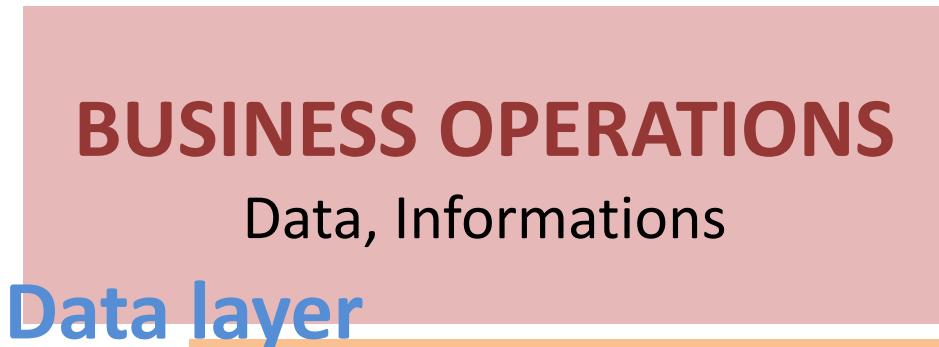
ŠTA JE BUSINESS INTELLIGENCE

BI

RELEVANTNE I POUZDANE INFORMACIJE

PRAVIM LJUDIMA
U PRAVO VРЕМЕ

Mesto Business Intelligence u poslovanju



Data layer

Core Business
IT Systems
Row Data



Data warehouse

Analyse the past ...
... to improve the future



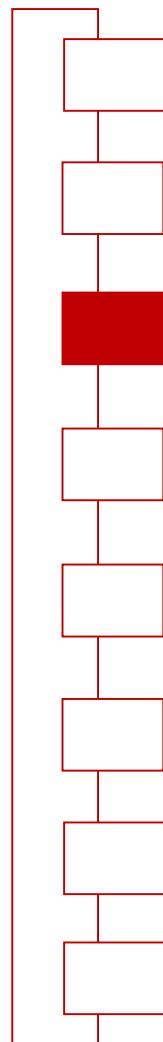
Data science

More Profit
Predictive
Mining

BUSINESS INTELLIGENCE obuhvata:

- ***Tehnike prikupljanja i integracije podataka***
(Extracting, Transformation, Loading – ETL i Data integration)
- ***Obrada i skladištenje podataka***
(Data Management i Data Warehouse - DW)
- ***Sintezu i analizu podataka***
(On Line Analytical Processing – OLAP)
- ***Prezentaciju podataka korisniku informacija***
(Information delivery)

Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

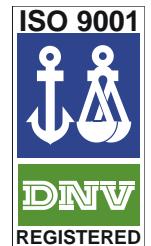
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



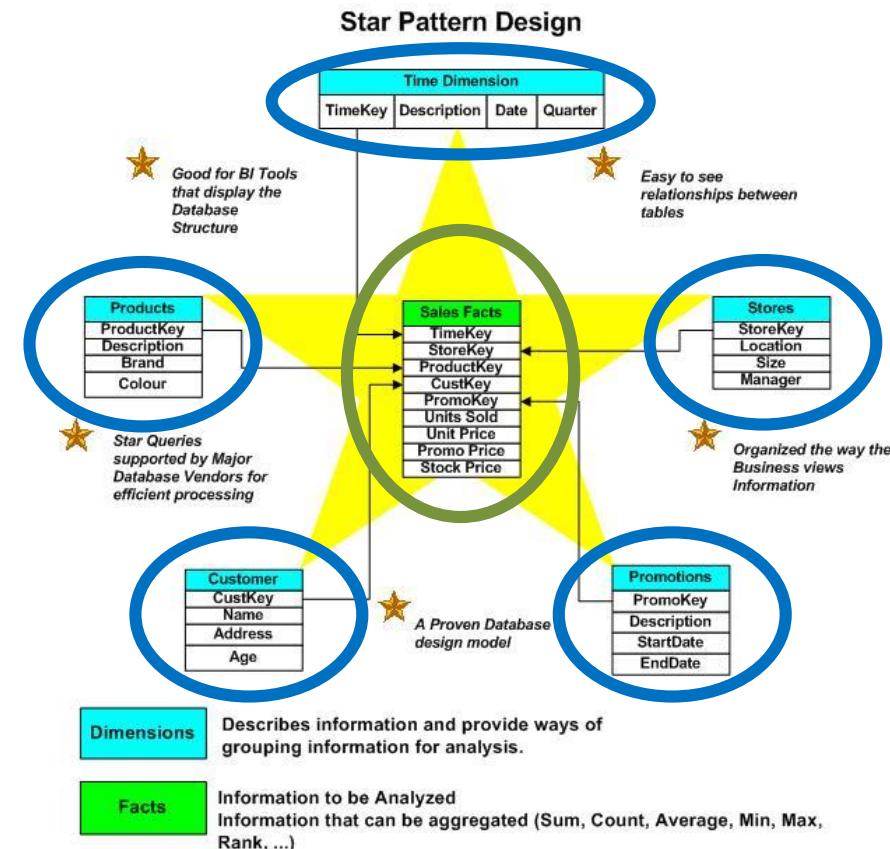
Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



ORACLE® Platinum Partner

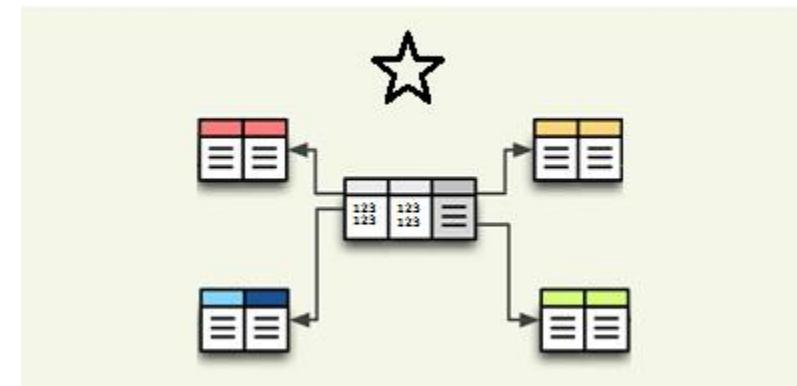
BI model se zasniva na:

- **činjenične tabele** - svi brojčani i kvantitavni podaci
- **dimenzione tabele** - opisni podaci za svaku oblast



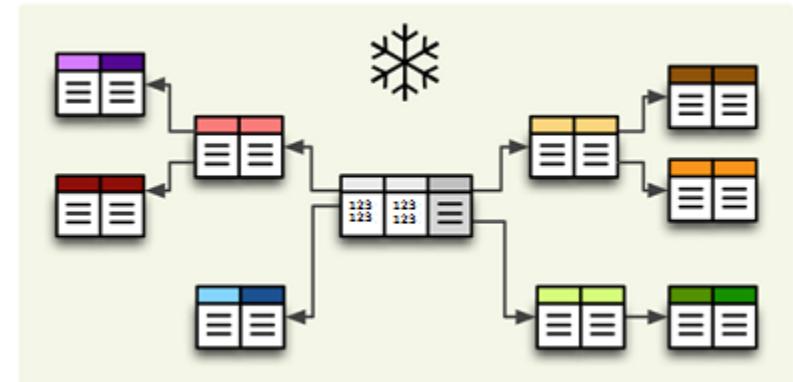
BI model se zasniva na:

- **činjenične tabele** - svi brojčani i kvantitativni podaci
 - **dimenzione tabele** - opisni podaci za svaku oblast
-
- **star – zvezdasta šeme**



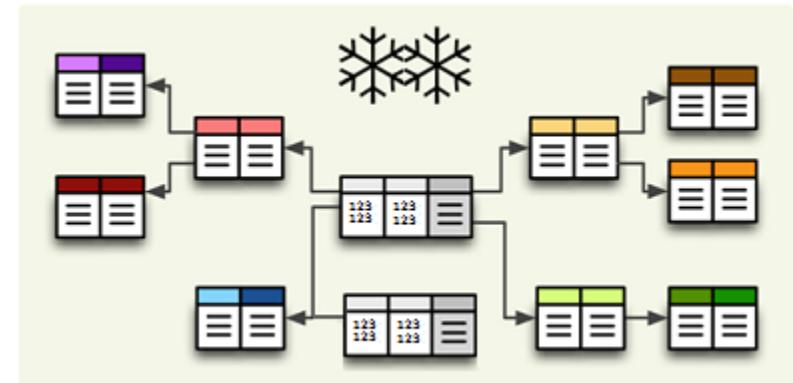
BI model se zasniva na:

- **činjenične tabele** - svi brojčani i kvantitavni podaci
 - **dimenzione tabele** - opisni podaci za svaku oblast
-
- star** – zvezdasta šeme
 - snowflake** – pahuljičasta šeme

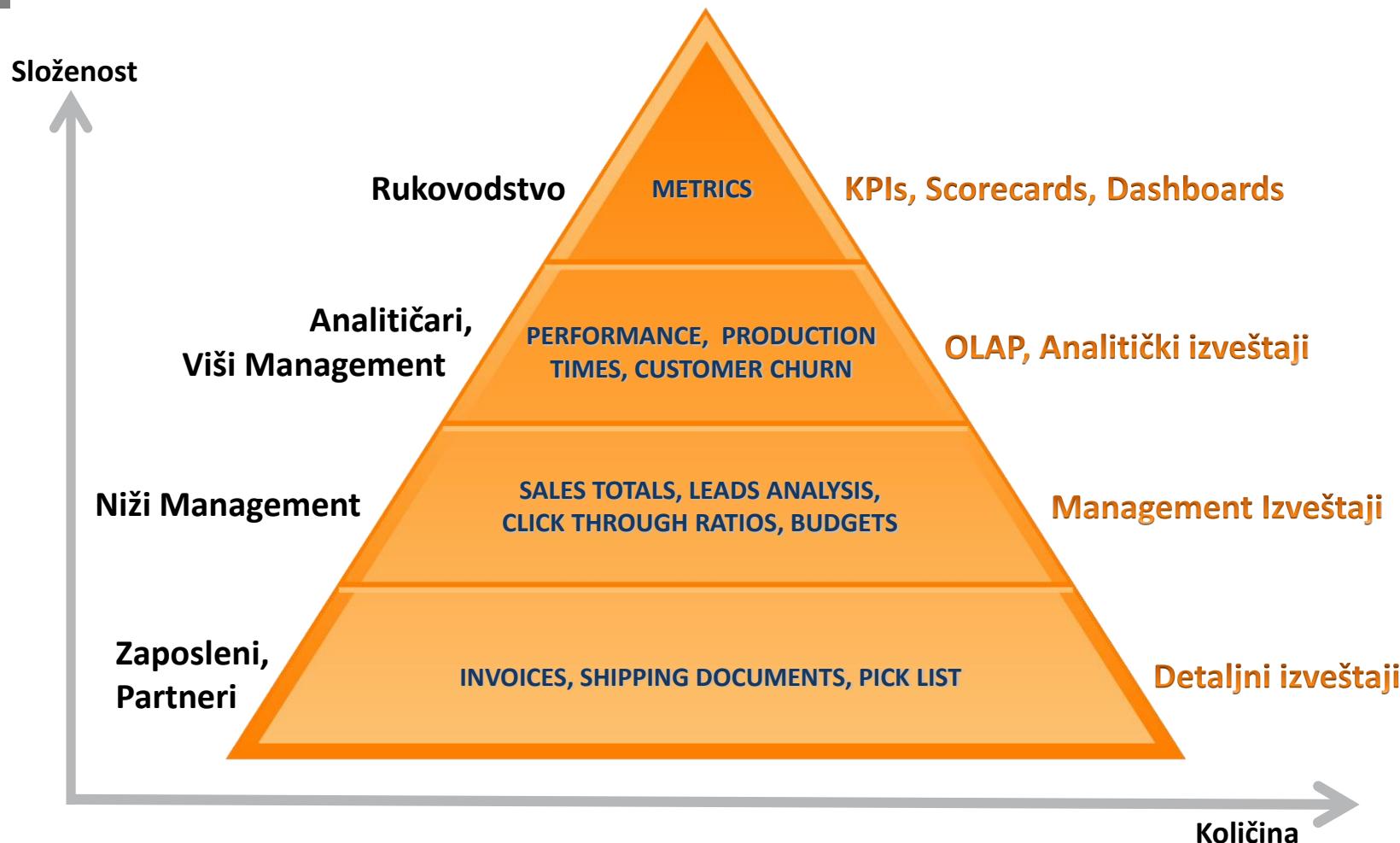


BI model se zasniva na:

- **činjenične tabele** - svi brojčani i kvantitativni podaci
 - **dimenzione tabele** - opisni podaci za svaku oblast
-
- star** – zvezdasta šeme
 - snowflake** – pahuljičasta šeme
 - galaxy** – galaksija šeme



Šta je Business Intelligence i kome je namenjen



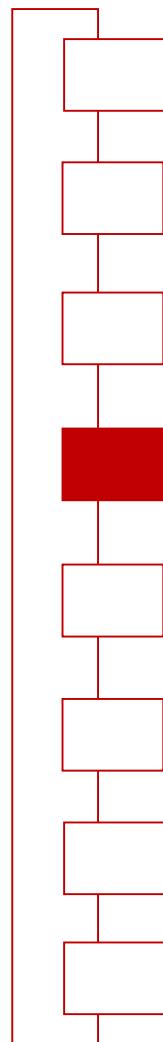
Karakteristike BI projekta:

- Nije klasičan IT projekat*
- Namenjen je menadžerima i donosiocima odluka, a ne izvršiocima poslovanja (komunikacija na “višem” nivou)*
- U izradi BI projekta učestvuje multidisciplinarni tim*
- Koristi se IT tehnikama, matematikom, statistikom, poslovnim znanjima*
- Koristi se skladište podataka, a ne baza podataka*
- Koriste se podaci stari jedan dan*
- Radi se sa analitičkim, ne sa transakcionim podacima*
- Pamte se svi podaci kroz vreme*
- Ne radi se za unapred poznato poslovanje, već se u toku analize i izrade projekta nadograđuje i menja model*
- Ne radi se klasično izveštavanje već analitičko i cilj mu je povećanje profitabilnosti i poboljšavanje poslovnih procesa*

Aktivnosti na projektu:

- Priprema projekta i snimak stanja, formiranje projektnog tima
- Instalacija i konfigurisanje sistema
- Gap analiza, definisanje poslovnih područja i izvora podataka
- Dizajn sistema (ETL, DWH, Izveštajno-analitički modul)
- Implementacija sistema (repositorijumi, dashboradi, scorecardi)
- Edukacija i obuka korisnika
- Testiranje
- Dokumentovanje sistema

Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

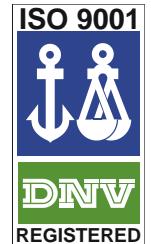
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health

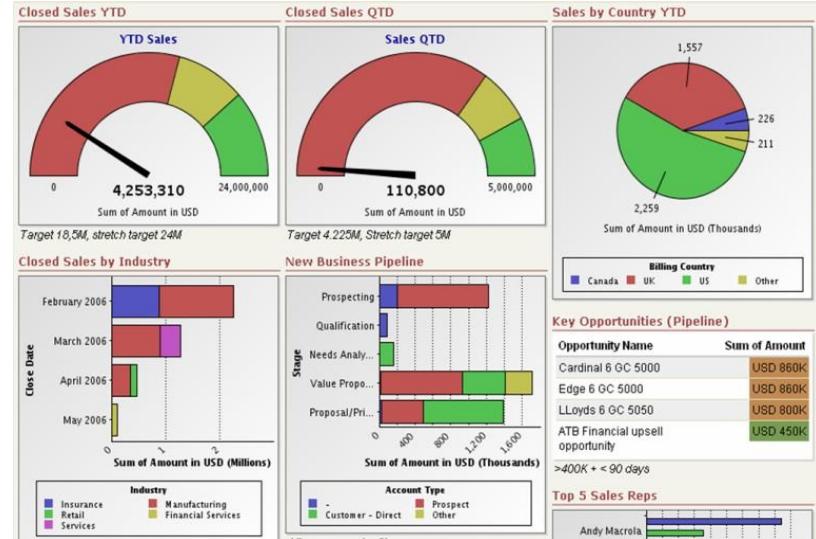


ORACLE® Platinum Partner

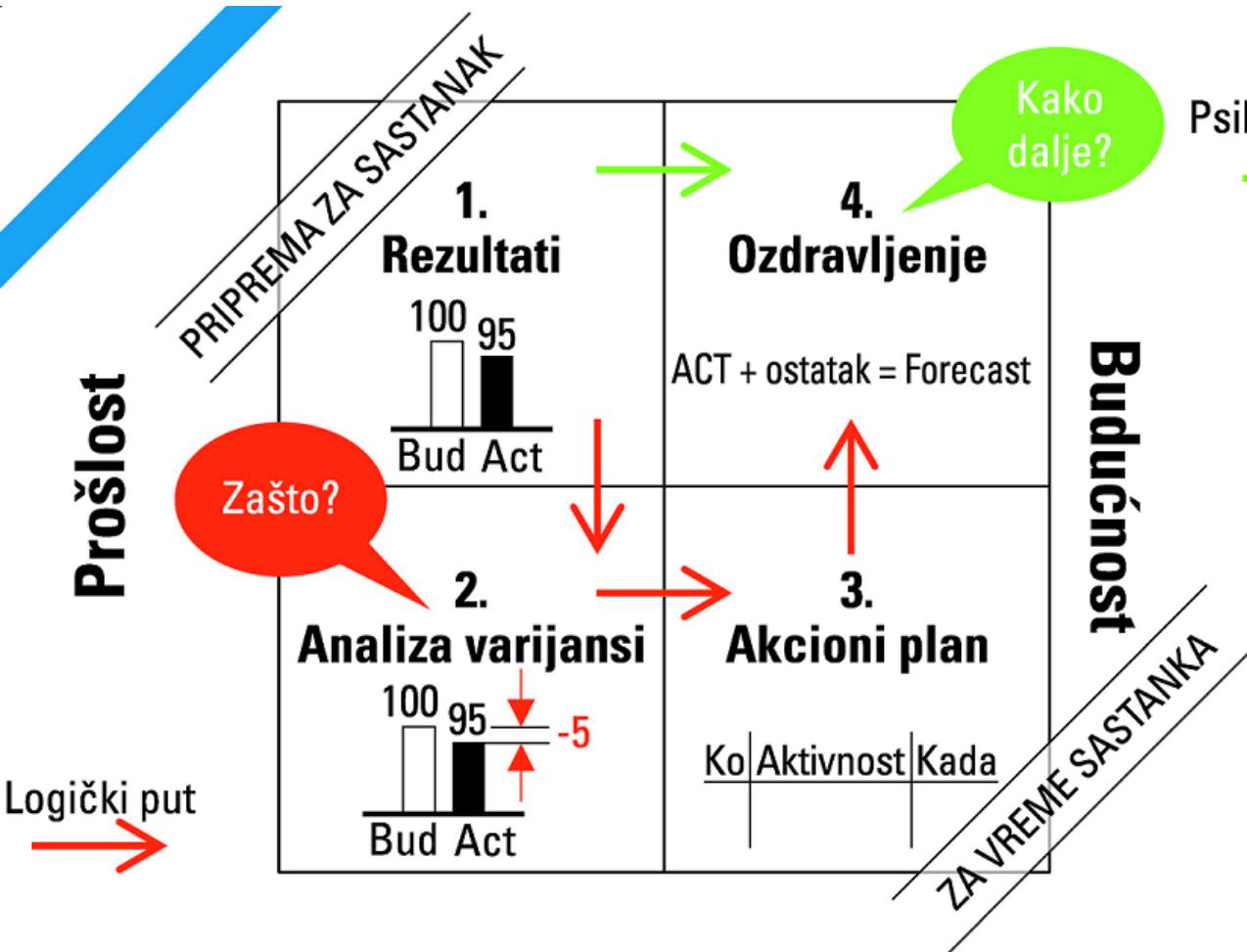
IZVEŠTAVANJE I ANALIZE U BI MODELU:

VRSTE IZVEŠTAJA I ANALIZA

- Redovni Izveštaji:
 - Eksterni (prema institucijama)
 - Interni (prema managementu)
 - Razne revizije
- Ad-hoc izveštaji i upiti
- Analitički izveštaji (OLAP)
- Drill-down analize, cross-sell analize
- Izrada analiza kojima se mogu planirati akcije i njihov učinak:
 - Vremenske serije - Time series
 - Trend analize
 - Prognoze - Forecast analize ...



BI4ERP – Struktura izveštaja i analiza



4 Windows form Analitička forma

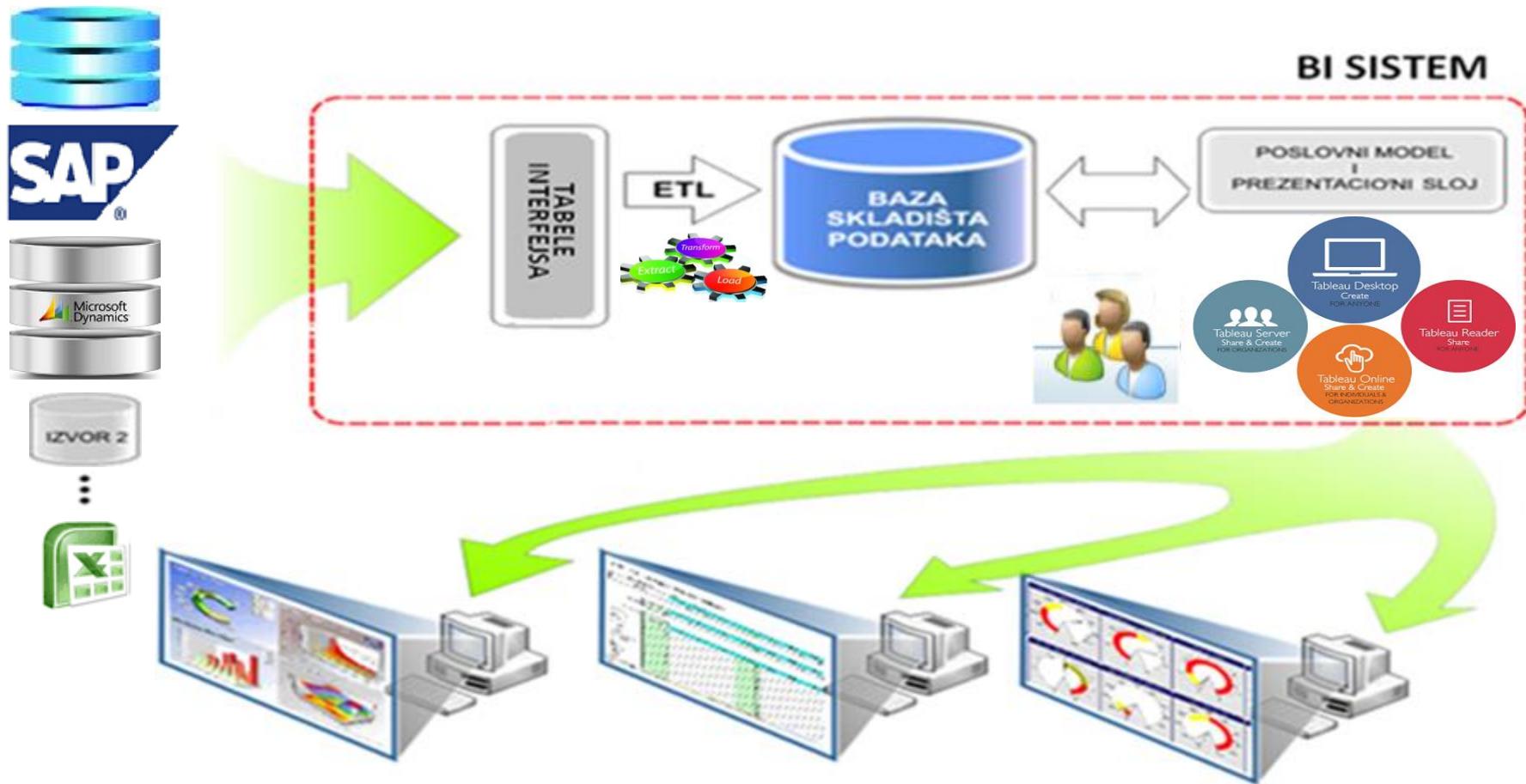
Psihološki put

- Rezultati**
Snimak stanja
- Analiza varijansi**
Analitički izveštaji
- Akcioni plan**
WhatIf analize
- Ozdravljenje**
Forecast, detalji

Tehnologija ponuđenog rešenja:



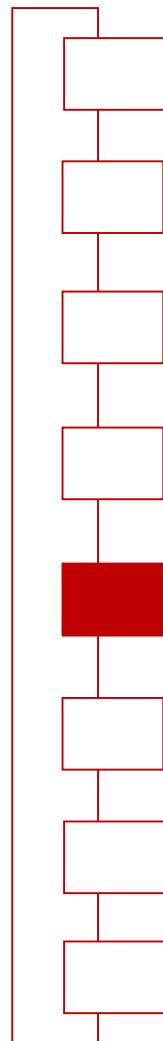
Predložena Arhitektura



IN2 BI4ERP

DEMO

Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

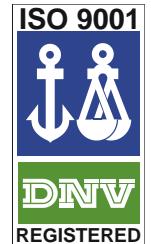
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



ORACLE® Platinum Partner

Data Analytics

- Prediktivna analiza
- Biznis analiza
- Biznis slučajevi
- Data Mining

expect
more.

Big Data

- Skladištenje velikih količina podataka
- Skladištenje različitih podataka
- Skladištenje velikom brzinom
- Analize podataka
- Tehnologije:

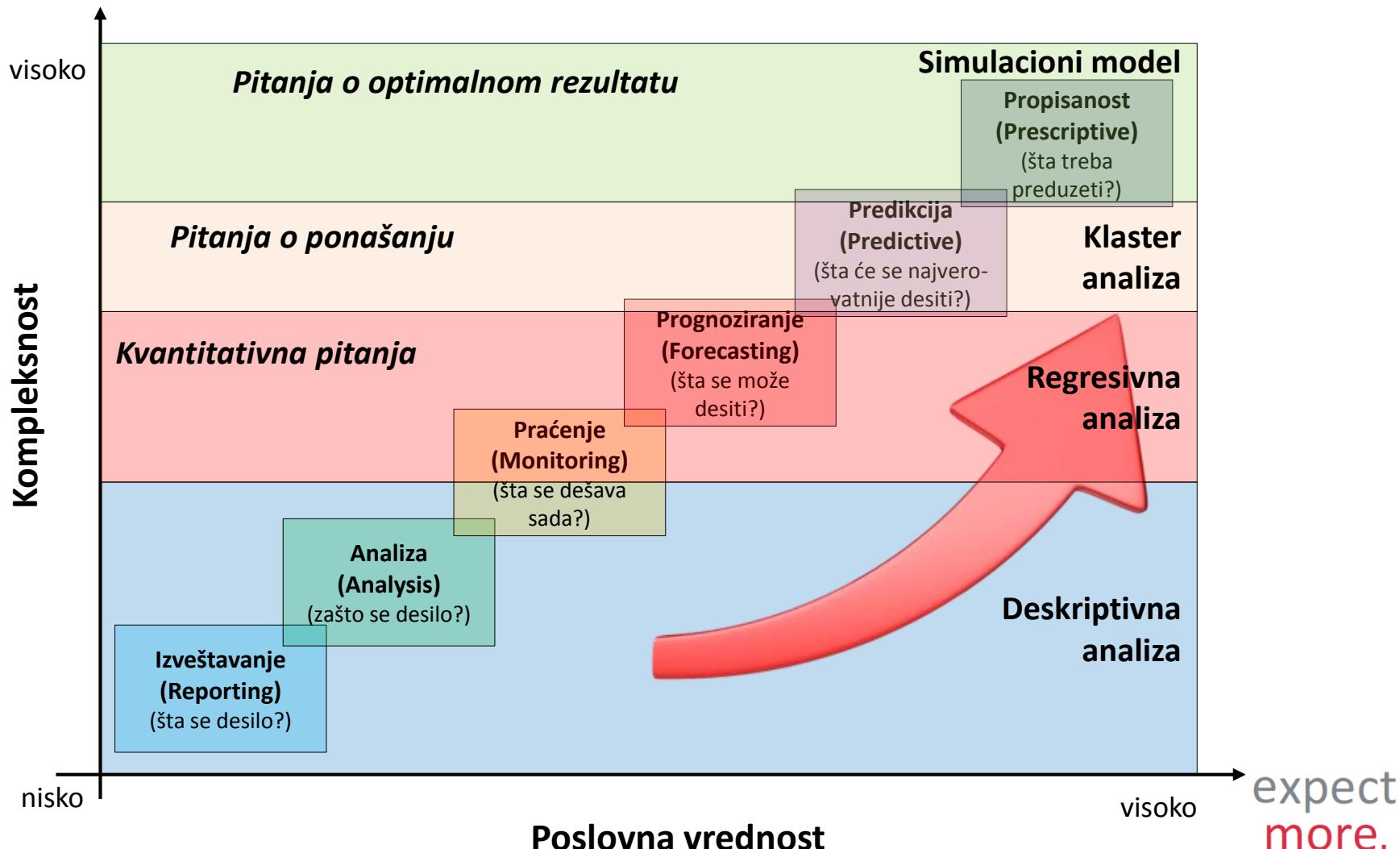


IN2data | Prediktivna analiza

IN2data Data **Science** Company:

- ❑ Multidisciplinarni tim **naučnika i profesionalaca** sa relevantnim **akademskim i poslovnim** znanjem, koji su značajni za isporuku naprednih analitičkih rešenja
- ❑ Projekti uključuju prediktivne modele i **big data** analitiku,
 - ✓ SNA – analizu socijalnih mreža,
 - ✓ segmentacione modele,
 - ✓ cross sell & up sell modele,
 - ✓ fraud detection modele,
 - ✓ churn predikciju,
 - ✓ scoring modele,
 - ✓ text mining modele,
 - ✓ sisteme za rano upozoravanje i
 - ✓ social CRM modele.

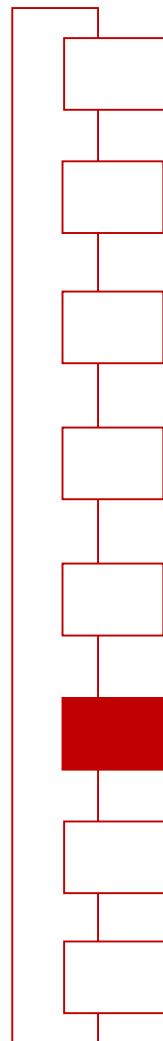
expect
more.



IN2data | Poslovni stubovi - Pillars



Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

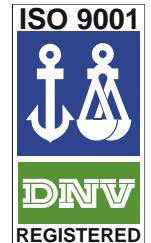
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



SNA analiza – Uvodne napomene

Analiza mreža

Analiza socijalnih mreža

- Podaci o mreži se predstavljaju u formi (društvenog) grafa
 - svaki **član** mreže predstavlja jedan **čvor** grafa
 - **relacije** među članovima mreže su **ivice** grafa
- Analiza mreže je primena skupa SNA metrika nad grafom

SNA analiza – karakteristike mreža

- ***Model mreže, zakonitosti mreže***

(Srazmerna i Nesrazmerna, “Mali svet”, Snaga slabih veza)

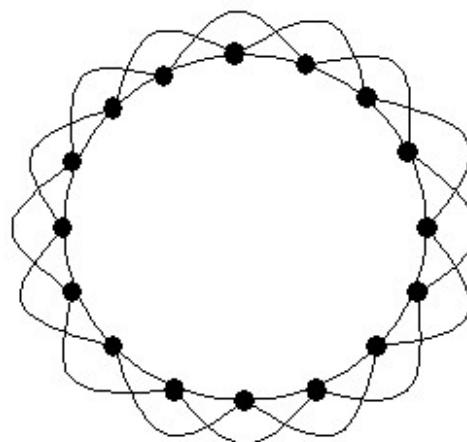
- ***Opis karakteristika mreže***

(Gustina, Povezanost, Struktuisanost)

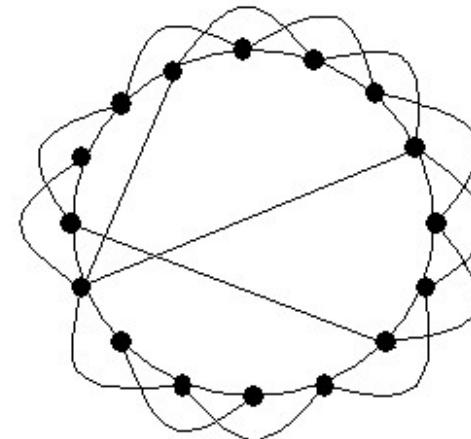
- ***Opis entiteta mreže***

(Zavisi od cilja analize mreže)

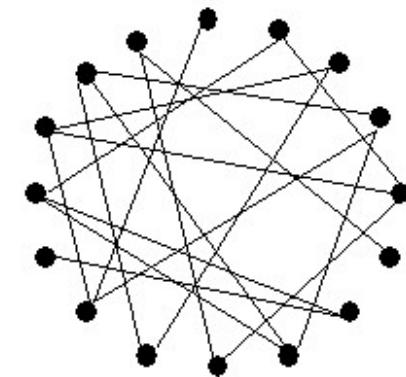
SNA analiza – Vrste socijalnih mreža



Regularna mreža



Mreža malog sveta



Slučajna mreža

0

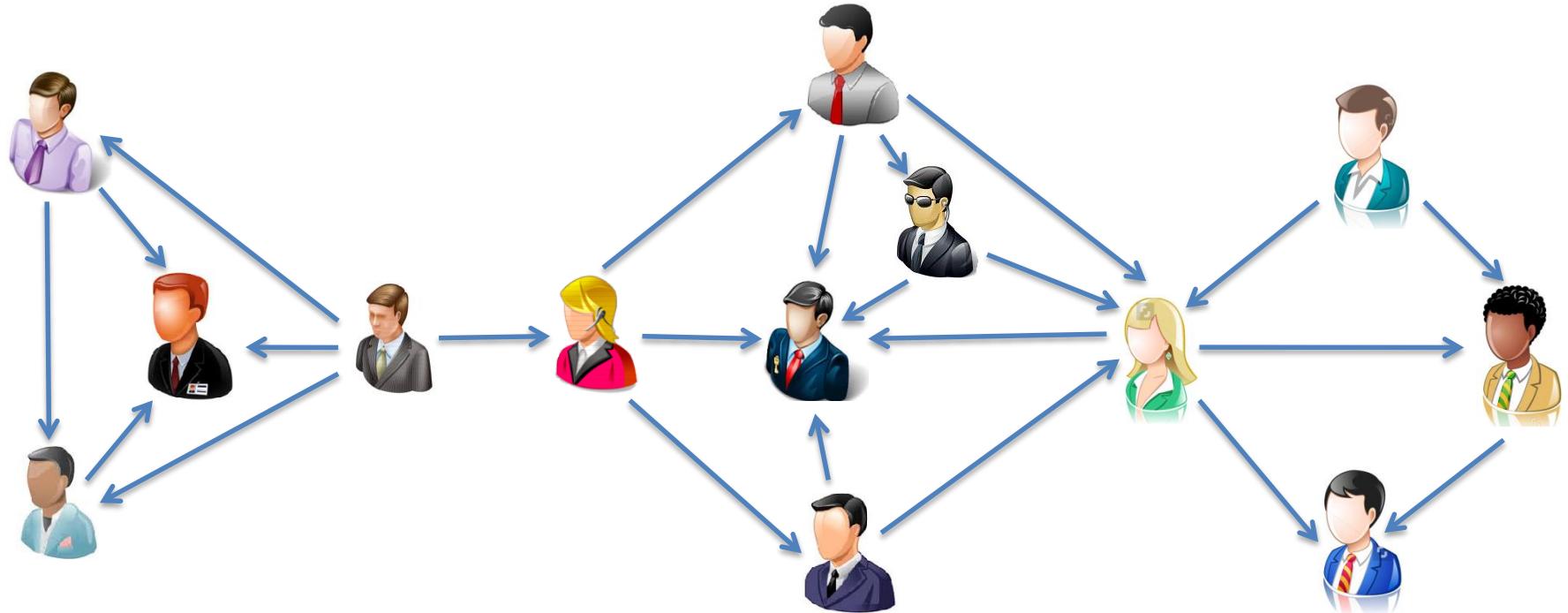
Verovatnoća prepovezivanja

1

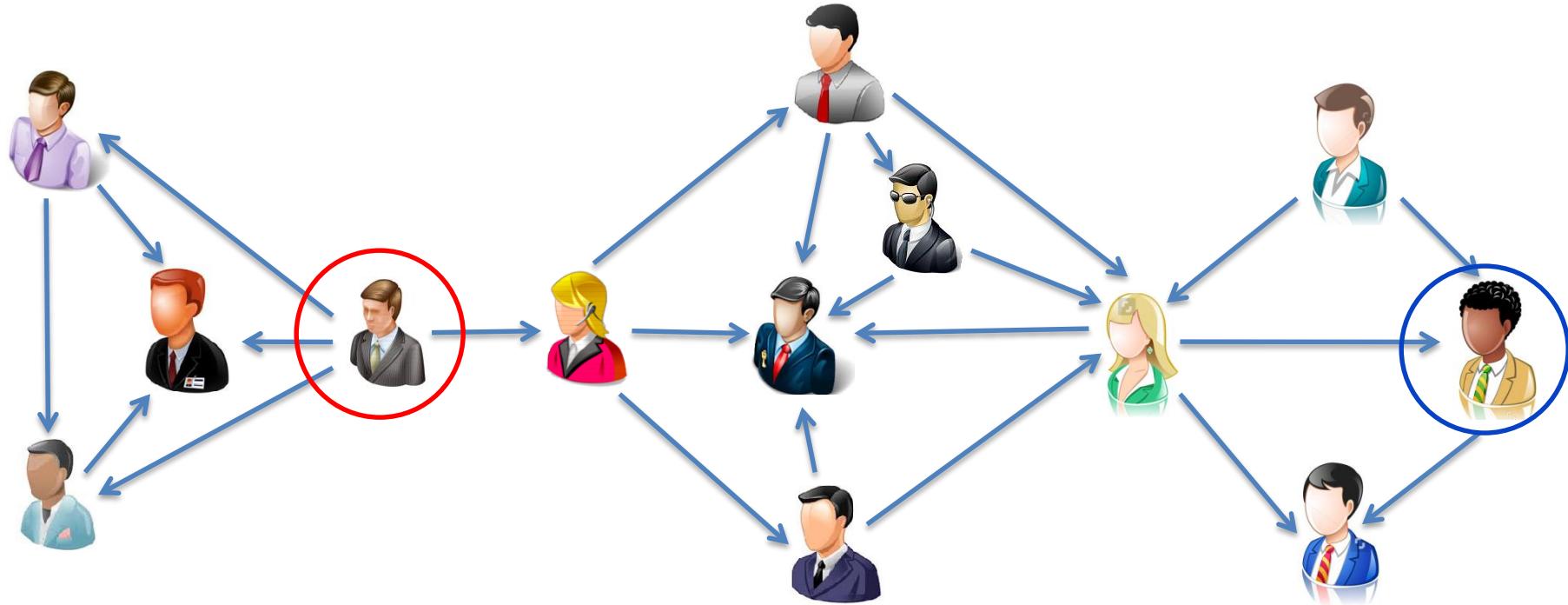
SNA analiza – SNA metrika

- Broj entiteta i veza**
- Broj komponenti** (Connected Components)
- Maksimalna geodetska distanca** (Maximum Geodesic Distance)
- Prosečna geodetska distanca** (Average Geodesic Distance)
- Gustina mreže** (Graph Density)
- Razvrstljivost/Asortativnost** (Assortativity)

SNA analiza – Praktična primena

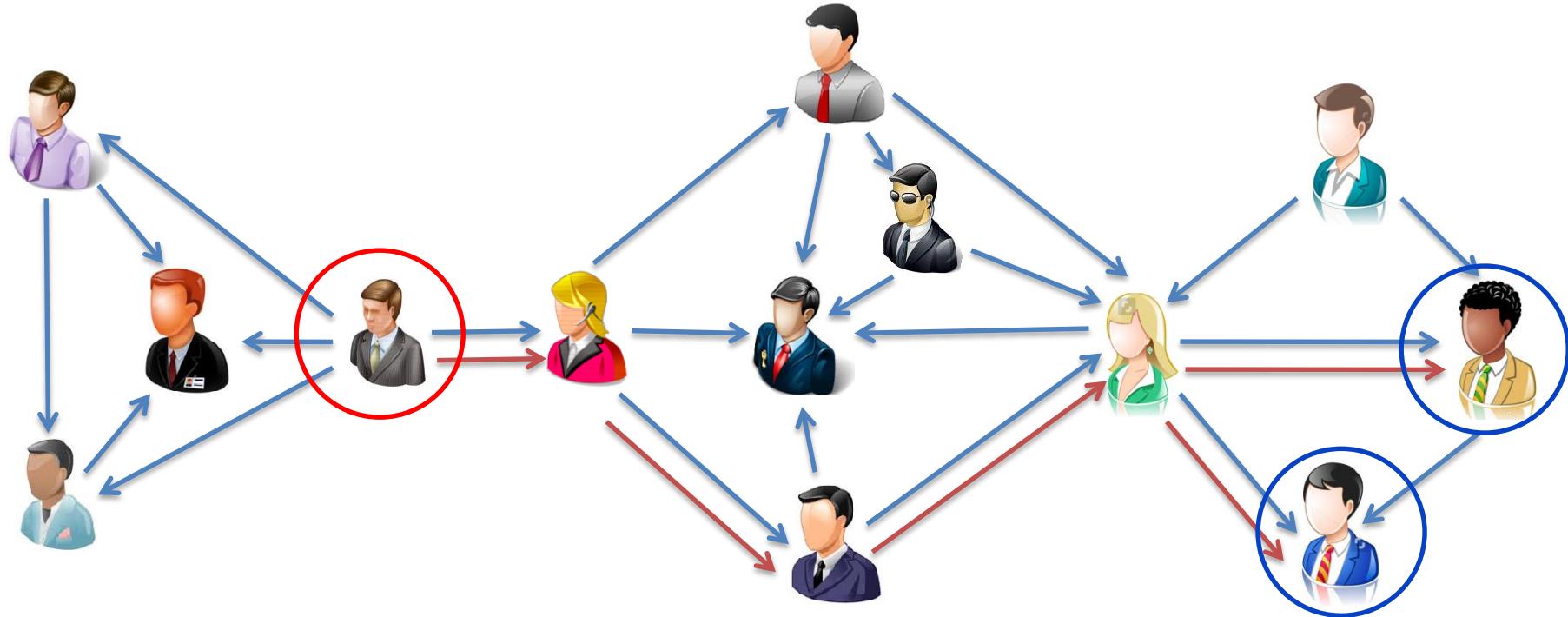


SNA analiza – Praktična primena



DIJAMETAR - Maksimalna geodetska distanca
(Maximum Geodesic Distance)

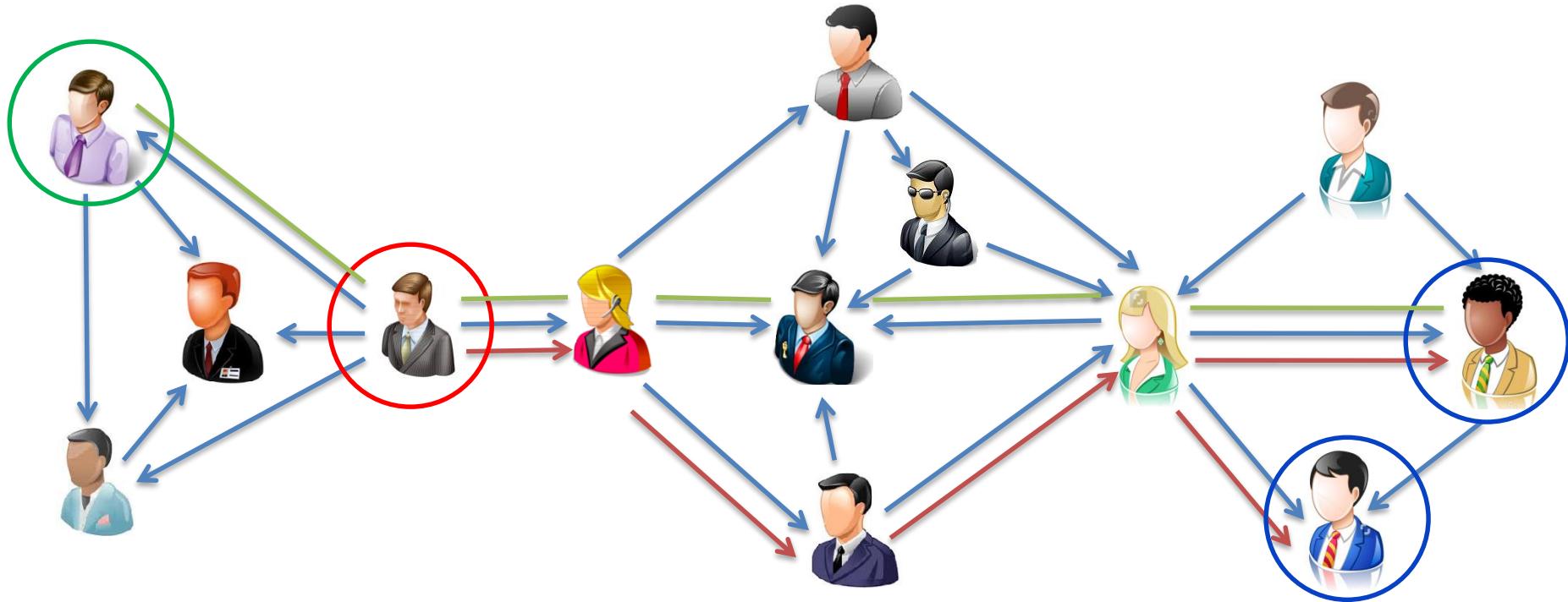
SNA analiza – Praktična primena



DIJAMETAR - Maksimalna geodetska distanca
(Maximum Geodesic Distance)

4

SNA analiza – Praktična primena

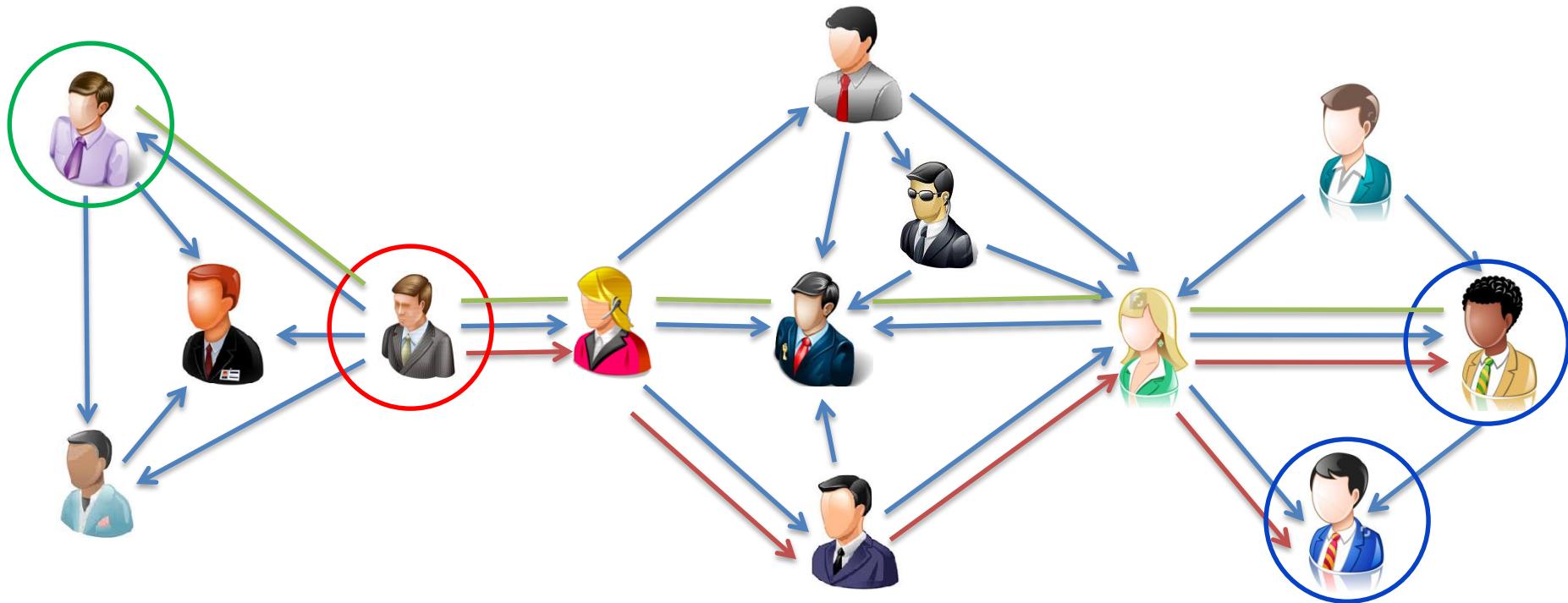


DIJAMETAR - Maksimalna geodetska distanca

(Maximum Geodesic Distance)

4 5

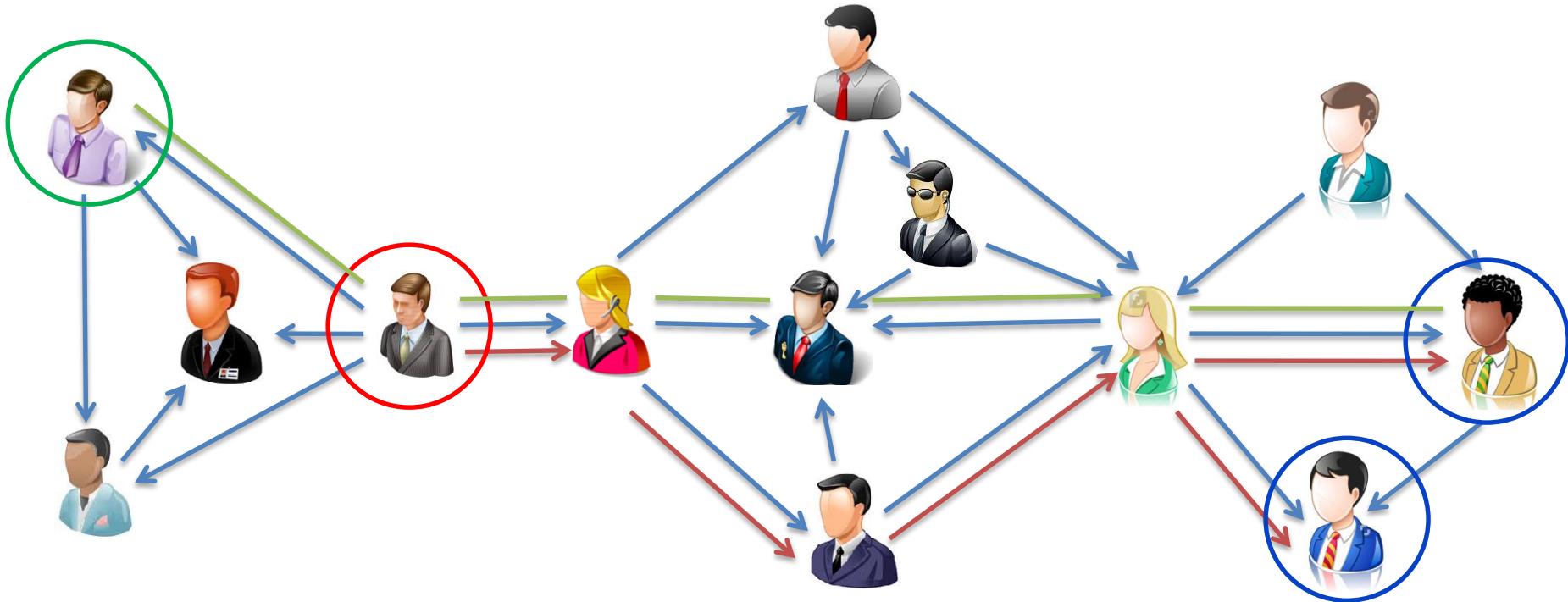
SNA analiza – Praktična primena



Prosečna geodetska distanca (Average Geodesic Distance)

2,27

SNA analiza – Praktična primena

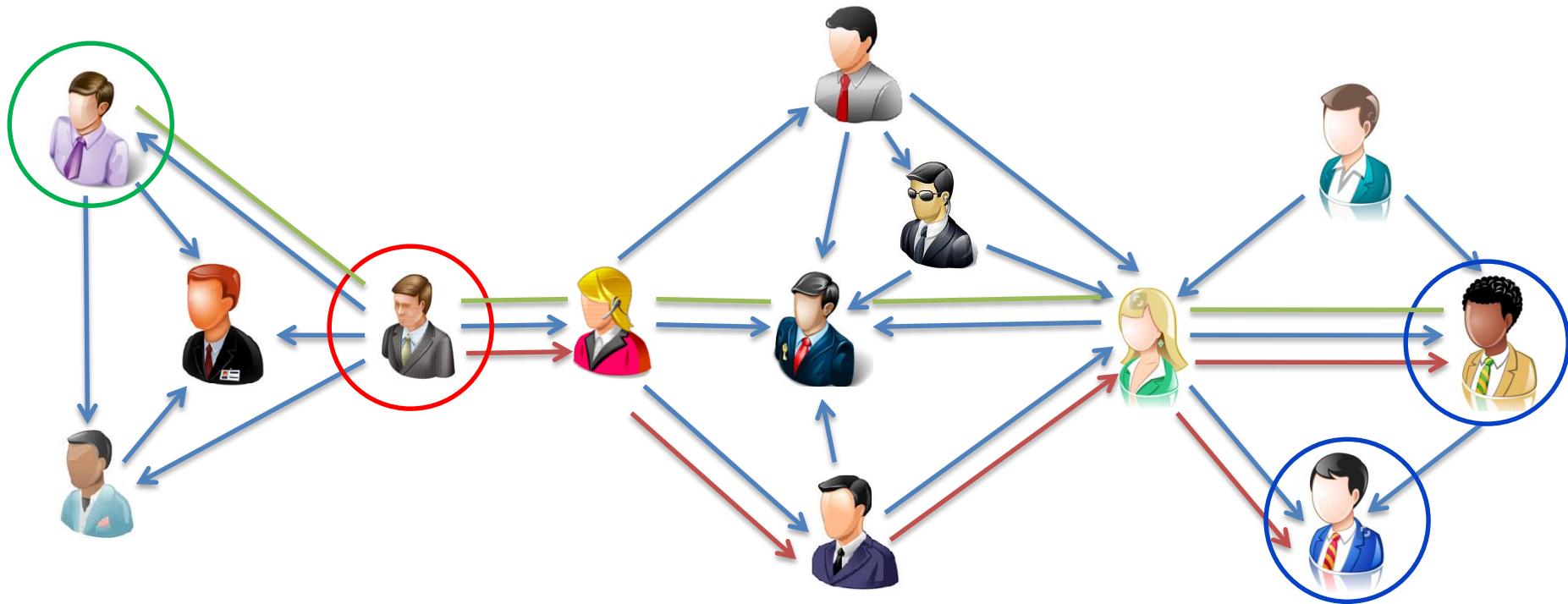


Gustina mreže
(Graph Density)

0,29

Broj svih mogućih veza
Broj ostvarenih veza

SNA analiza – Praktična primena



Asortativnost (Assortativity)

-1 do 1

Spajanje čvorova sa sličnim čvorovima

SNA analiza – Identifikacija ključnih osoba

in2

➤ MERE CENTRALITETA (Centrality)

- ❑ **Degree** (Broj drugih čvorova sa kojima je dati čvor povezan)

U slučaju usmerenog grafa, razlikujemo:

- **In-degree** (Broj ivica koje uviru u dati čvor)
 - **Out-degree** (Broj ivica koje polaze iz datog čvora)
 - **Total-degree** (Ukupan broj ivica koje polaze iz ili uviru u dati čvor)

Strukturni i Socijalni renome (Structural & Social Prestige)



SNA analiza – Identifikacija ključnih osoba

➤ *MERE CENTRALITETA (Centrality)*

- Degree
- Betweenness centrality (Identifikacija čvorova koji se pojavljuju kao posrednici u komunikaciji između drugih čvorova mreže)
Identifikacija čvorova posrednika i čvorova pucanja.



SNA analiza – Identifikacija ključnih osoba

➤ *MERE CENTRALITETA (Centrality)*

- Degree
- Betweenness centrality
- Closeness centrality (Izračunati prosečnu dužinu najkraćih putanja od čvora v do svih ostalih čvorova mreže)

Definiše brzinu kojom informacije, polazeći od nekog izabranog čvora mogu stići do ostalih čvorova mreže.

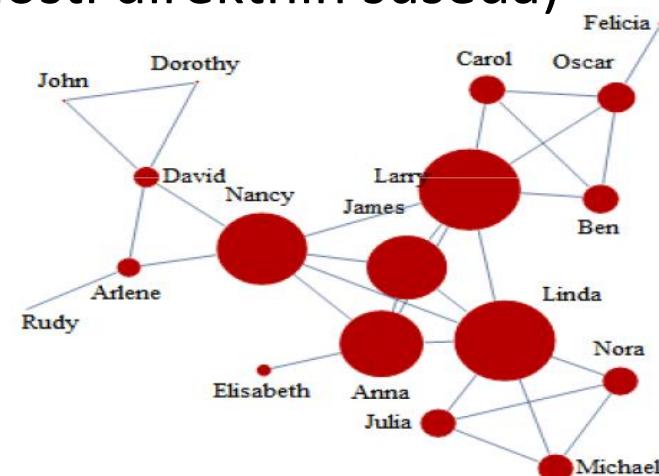


SNA analiza – Identifikacija ključnih osoba

in2

➤ **MERE CENTRALITETA (Centrality)**

- Degree**
 - Betweenness centrality**
 - Closeness centrality**
 - Eigenvector centrality** (Određuje ko je u mreži povezan sa uticajnim članovima mreže – suma vrednosti direktnih suseda)
 - Identifikacija lidera, tj. uticajnih članova mreže.
 - “Page rank” algoritam.



SNA analiza – Metrika centralnosti

INTERPRETACIJA METRIKA CENTRALNOSTI U KONTEKSTU DRUŠTVENIH MREŽA

Degree

Koliko drugih ljudi data osoba može direktno kontaktirati?

Betweenness

Koliko je verovatno da će data osoba imati ulogu posrednika između nekih drugih dvoje ljudi iz mreže ?

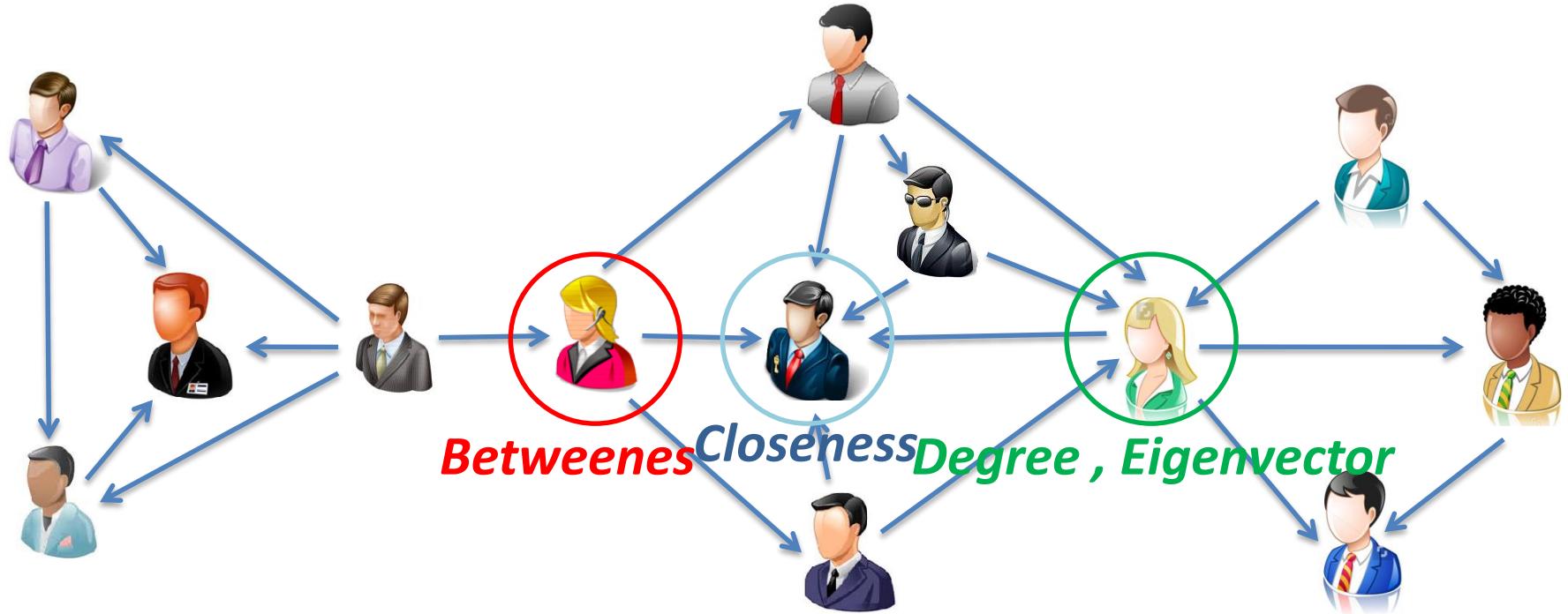
Closeness

Koliko brzo data osoba može dopreti do svake druge osobe u mreži? mreže?

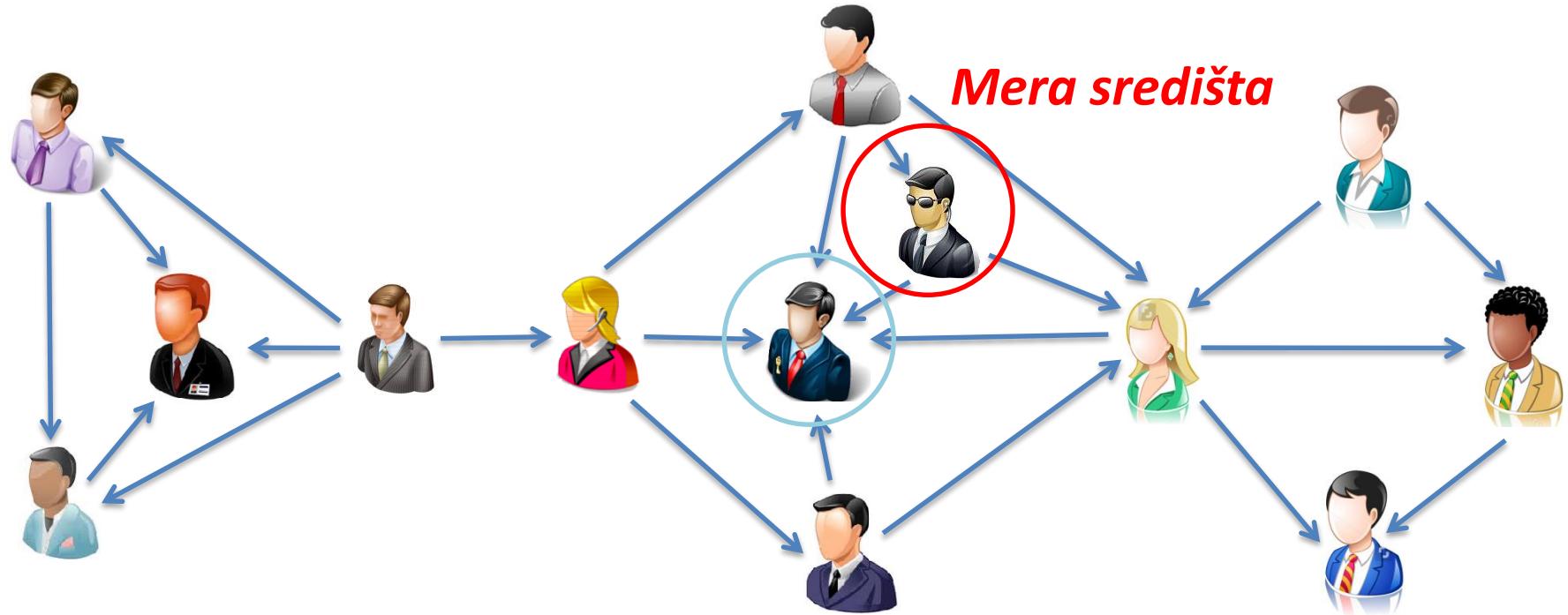
Eigenvector

U kojoj meri je data osoba povezana sa uticajnim tj. dobro povezanim članovima mreže?

SNA analiza – Praktična primena



SNA analiza – Praktična primena



Mera središta – Vezan za čvorove sa dobrim autoritetom

SNA analiza – Softverski alati za analizu

- **Gephi** (podrška za sve platforme, besplatan, open source)
- **R paketi za SNA** (podrška za sve platforme, besplatan, open source)
- **Pajek** (Windows, besplatan)
- **NodeXL** (Windows, besplatan, open source; integriše se u Excel)
- **UCInet** (Windows, komercijalni)



R Programming.net



UCINET Software

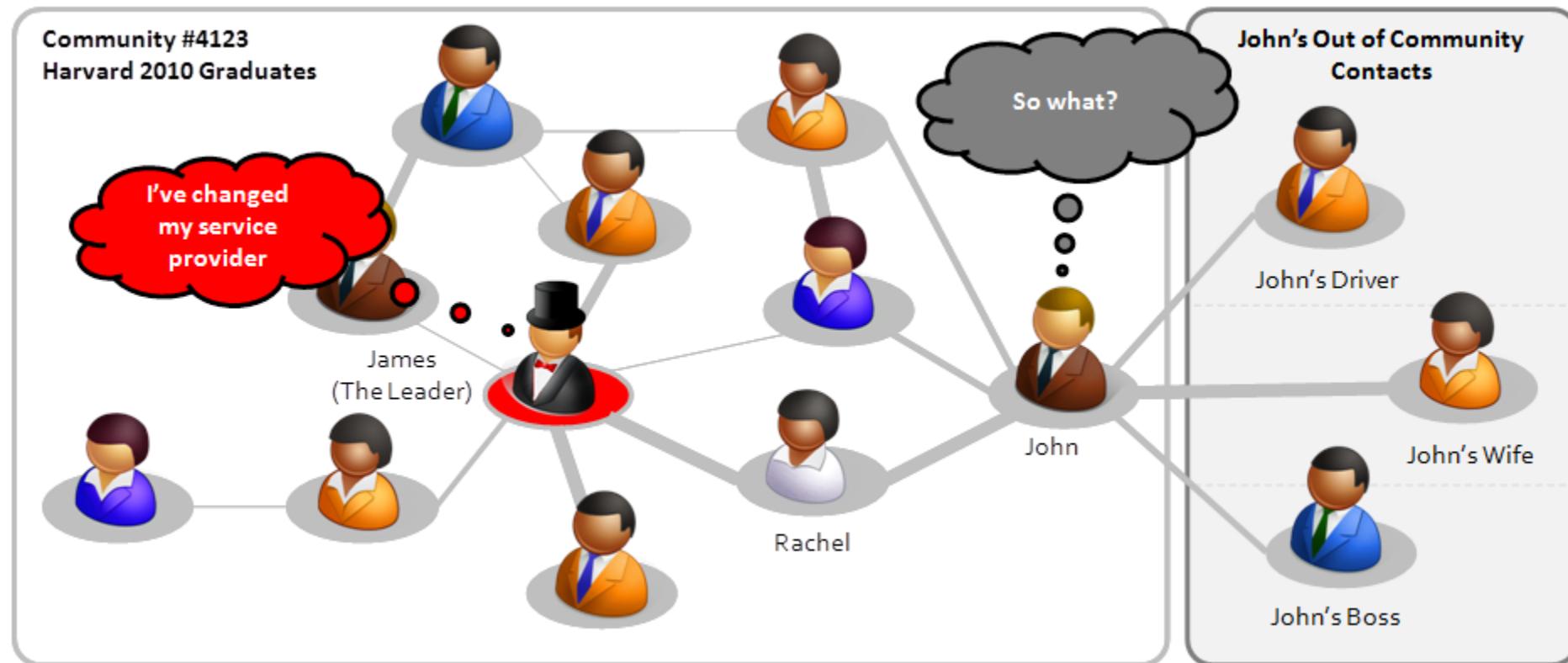


Pajek



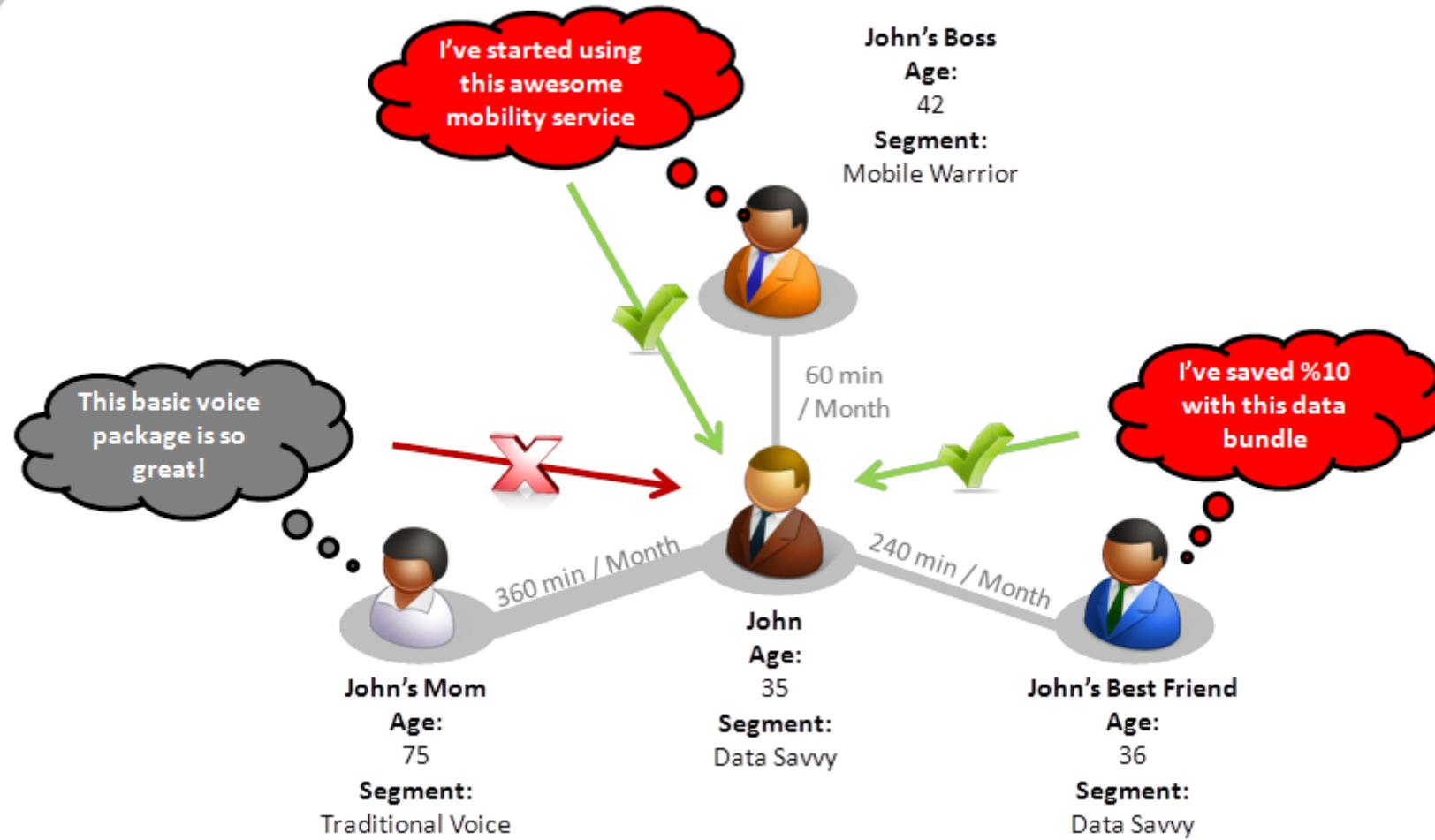
Gephi
makes graphs **handy**

SNA analiza – Praktična primena

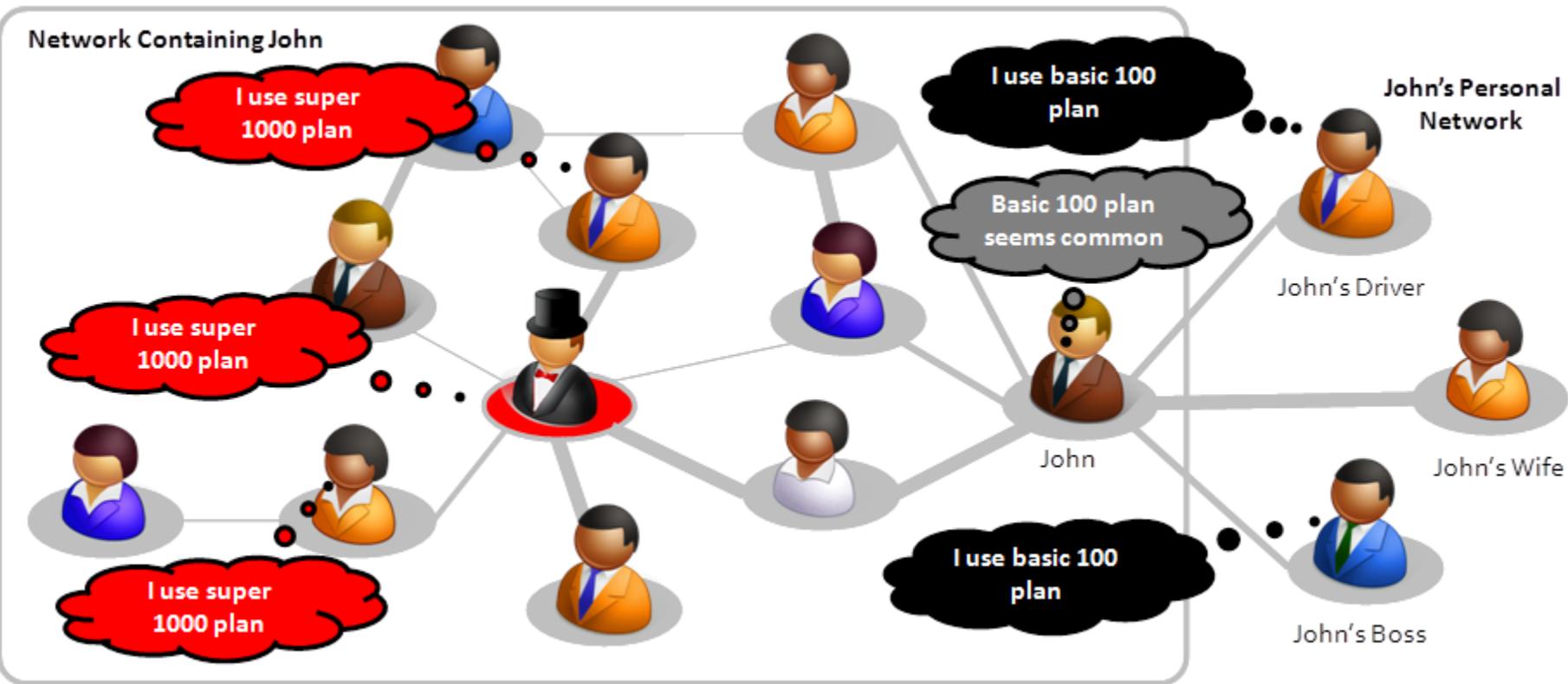


Pitanje – Da li lideri imaju više uticaja na ostale članove mreže?

SNA analiza – Praktična primena

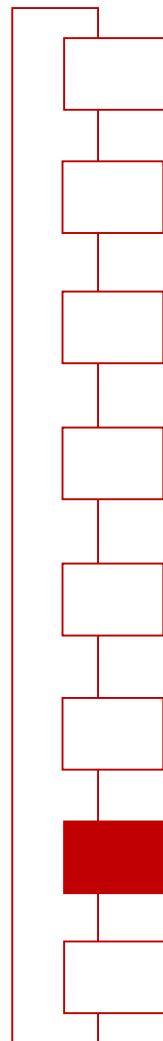


SNA analiza – Praktična primena



Pitanje – Da li na pojedinca najviše utiče njegova primarna mreža?

Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

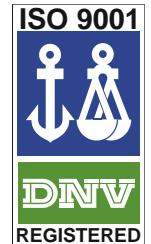
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



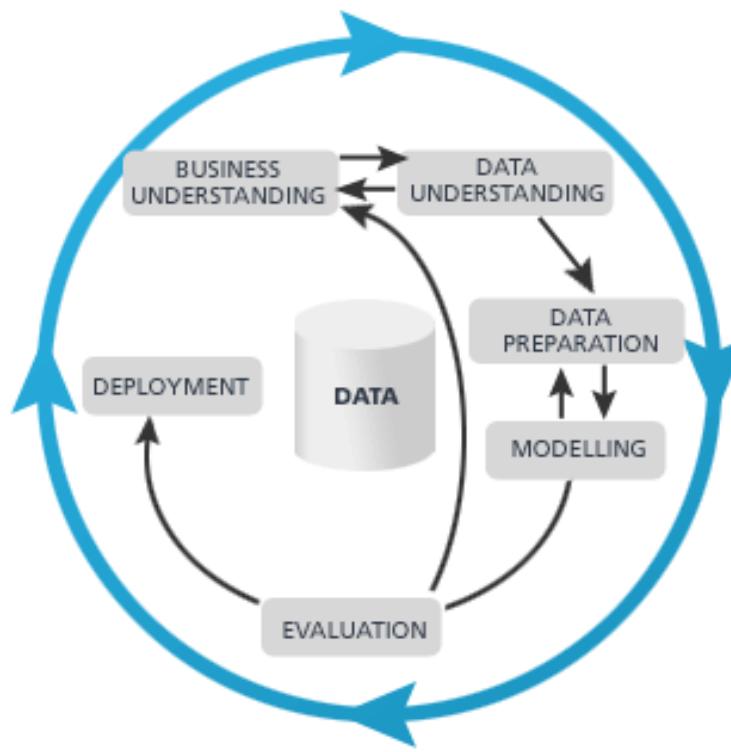
Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



ORACLE® Platinum Partner

Primena istraživanja i analize podataka u cilju formiranja namenske polise osiguranja za novog klijenta osiguravajuće kuće

- ❑ Cross-Industry Standard Process for Data Mining (CRISP-DM)



Opis atributa

- Svaki od 5683 klijenta je opisan sa 11 atributa (4 numerička i 7 kategoričkih)

R. br.	Naziv atributa	Značenje atributa
1	Godine	Godište klijenta
2	Pol	Pol klijenta
3	Zanimanje	Zanimanje klijenta
4	Tezina	Težina klijenta
5	Visina	Visina klijenta
6	Adresa	Adresa stanovanja klijenta
7	Grad	Mesto stanovanja klijenta
8	Pristupna_dob	Sa koliko godina je klijent potpisao ugovor, polisu osiguranja, sa osiguravajućom kućom
9	Trajanje_osiguranja	<i>Period na koji je potписан ugovor klijenta sa osiguravajućom kućom</i>
10	Premija	<i>Iznos premije koju korisnik plaća</i>
11	Dinamika	<i>Način na koji klijent izvršava svoje plaćanje</i>

Modelovanje i evaluacija

- Ulazni skup podataka je podeljen na *train* i *test* skupove u odnosu 80% naspram 20% u korist *train* skupa
- ***Explorative Data Analysis (EDA)***
 - Problem nedostajućih vrednosti
 - Problem koreliranih atributa
 - Transformacija atributa (*ln_visina*, *ln_tezina*, *ln_godine_klijenta*, *sqrt_pristupna_dob*, *flag_visina*, *flag_tezina*)
 - Istraživanje odnosa atributa klase i ostalih atributa (matrica korelacija sa za svaki ciljni atribut)
 - Istraživanje nominalnih atributa
 - Istraživanje numeričkih atributa
 - Analiza veza između više atributa i ciljnih atributa

Modelovanje i evaluacija

□ Deskriptivne metode

- **Klasterovanje** - "proces organizovanja objekata u grupe čiji članovi su na neki način slični"
- **Analiza asocijativnih pravila** - postupak pri kojem se identifikuju elementi koji se pojavljuju zajedno u nekom događaju

□ Prediktivne metode

- **Stabla odlučivanja (DTREE)** - omogućuju hijerarhijski način prikaza znanja. Listovi u stablu odlučivanja prikazuju sva moguća rešenja zadatog problema. Listovi se nazivaju i „čvorovima odgovora“, dok se svi ostali čvorovi u stablu nazivaju „čvorovima odlučivanja“. Čvorovi odlučivanja prikazuju pitanja ili odluke.
- **k-najbližih suseda (KNN)** - osnovna ideja ove vrste klasifikacije jeste da se na osnovu prepoznatih klasa, prema određenim specifičnostima definisanih u okviru obučavajućeg skupa, novi testni podatak može svrstati u neku od postojećih klasa na osnovu svojih karakteristika.
- **Naivni Bayes (NB)** - predstavlja metodu za klasifikaciju koja se oslanja na teoriju verovatnoće tj. na Bayes-ovu teoremu. Koristi se u situacijama kada obučavajući skup sadrži greške ili kada postoje još neki dodatni, nama nepoznati faktori, koji utiču na klasifikaciju.

Model i evaluacija

□ Rezultati evaluacije

		Trajanje osiguranja	Premija	Dinamika
Stabla odlučivanja		Model sa kriterijumom gini_index	Gain ratio sa selekcijom	Accuracy sa selekcijom
	tačnost	58.64%	86.12%	93.93%
	profit	1,191,950KM	1,882,400KM	2,096,700KM
K-najbližih suseda		K=3	K=8 sa selekcijom	K=8 sa selekcijom
	tačnost	56.53%	86.03%	94.72%
	profit	1,482,250KM	1,886,920KM	2,144,550KM
Naivni Bayes		Bez selekcije	Sa selekcijom	Sa selekcijom
	tačnost	61.48%	87.53%	94.26%
	profit	1,249,500KM	1,928,150KM	2,106,800KM

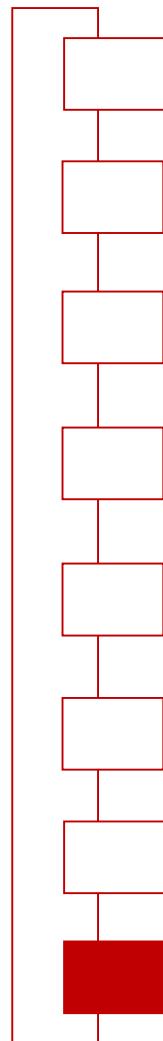
Model i evaluacija – Zaključci I

- *trajanje_osiguranja* - formirani modeli nisu dovoljno dobri za realnu upotrebu.
 - U interakciji sa zaposlenima osiguravajuće kuće čiji podaci su korišćeni u ovom radu, zaključeno je da je period na koji će klijent potpisati polisu osiguranja jako teško predvideti.
- *premija* - tačnost je relativno visoka - 87.53%, a preciznost i odziv za klasu od najvećeg interesa (premija manja od 50KM) uzimaju vrednosti 90.37% odnosno 97.67%.
 - Ovi rezultati vode do zaključka da je ulazni set podataka, naročito atributi *adresa* i *dinamika*, dovoljan da se formira zadovoljavajuće i u praksi upotrebljiv model za predikciju iznosa premije koju klijent treba da plaća.

Model i evaluacija – Zaključci II

- *dinamika* - tačnost iznosi 94.72%, a za preciznost i odziv za klasu od najvećeg interesa (mesečno plaćanje) uzimaju vrednosti 98.14% odnosno 99.28%.
 - Ovi rezultati vode do zaključka da je ulazni set podataka sasvim dovoljan za formiranje kvalitetnog modela za predikciju dinamike plaćanja premija. Naročito su bitni atributi: *tezina*, *visina*, *pristupna dob*, *adresa*, *trajanje osiguranja* i *premija* koji su primenom selekcije primećeni kao bitni za formiranje modela.

Agenda:



O IN2 Grupi

Šta je Business Intelligence i kome je namenjen

BI modeli i karakteristike BI projekata u praksi

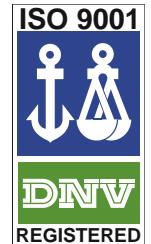
Demonstracija BI4ERP rešenja na Tableau platformi

In2data – Data Science company

Primena SNA u praksi

Praktična primena Data Mininga u Osiguranju

Pitanja i diskusija



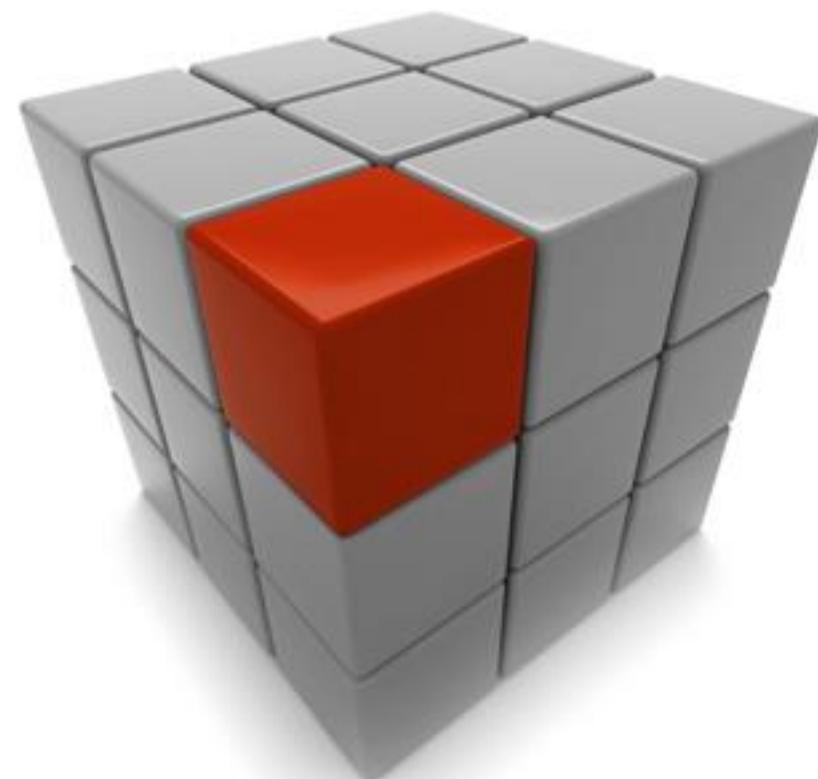
Microsoft Partner
2013 Partner of the Year Winner
Public Sector: Health



ORACLE® Platinum Partner



we think the way you need IT



IN2 d.o.o.

Milutina Milankovića 9ž/VI

11070 Beograd, SRB

tel: +381 11 2123 023

www.in2.rs