

12.07.2012. године
Београд,

УНИВЕРЗИТЕТ У БЕОГРАДУ
Веће научних области техничких
наука

Београд
Студентски трг бр 1

У прилогу дописа достављамо потребну документацију за давање сагласности на предлог теме докторске дисертације кандидата Драгољуба (Урош) Крнете.

Захтев се доставља у складу са чланом 54 Закона о Универзитету.

С поштовањем,

СЕКРЕТАР ФАКУЛТЕТА

Гордана Алексић Славински, дипл. Правник

Факултет организационих наука

УНИВЕРЗИТЕТ У БЕОГРАДУ

04-03-11/94
(Број захтева)

Веће научних области техничких наука
(Назив стручног већа коме се захтев упућује
сходночл 6. и члану 7. ст. 1. овог Правилника)

12.07.2012.
(Датум)

ЗАХТЕВ
за давање сагласности на предлог теме докторске дисертације

Молимо да, сходо члану 57. ст. 3. Закона о универзитету («Службени гласник РС» бр. 21/02), дате сагласност на предлог теме докторске дисертације: (пун назив предложене теме докторске дисертације)
Аутоматизација физичког пројектовања складишта података проширеног Data Vault приступа

НАУЧНА ОБЛАСТ: ОРГАНИЗАЦИОНЕ НАУКЕ

ПОДАЦИ О КАНДИДАТУ:

1. Име, име једног од родитеља и презиме кандидата:
Драгољуб (Урош) Крнета
2. Назив и седиште Факултета на коме је стекао високо образовање: Технички факултет "Михајло Пупин"
3. Година дипломирања: 2006. година
4. Назив приступног рада кандидата: Аутоматизација физичког пројектовања складишта података проширеног Data Vault приступа
5. Назив факултета на коме је приступни рад одбрањен: ФОН
6. Година одбране приступног рада: 2012. година
7. Назив факултета на коме је кандидат завршио докторске студије: /
Одсек, смер или група: /
Година завршетка докторских студија: _____

Обавештавамо вас да је: НАСТАВНО НАУЧНО ВЕЋЕ ФОН-А
(назив надлежног тела факултета)

на седници одржаној: 11.07.2012. године размотрило предложену тему и закључило да је тема подобна за израду докторске дисертације.

ДЕКАН ФАКУЛТЕТА

Прилог:

1. Предлог теме докторске дисертације са образложењем.
2. Акт надлежног тела факултета о подобности теме за израду докторске дисертације.

Проф. др Милан Мартић



Република Србија
Универзитет у Београду
Факултет организационих наука

Број индекса: 3/2009

04-03-11/94

Датум: 12.07.2012.

На основу члана 161. Закона о општем управном поступку и службене евиденције, издаје се

УВЕРЕЊЕ О ПОЛОЖЕНИМ ИСПИТИМА

Драгољуб Крнета, име једног родитеља Урош, рођен 06.01.1966.године, Зебе, Босна и Херцеговина, 2009/10. школске године уписан на докторске академске студије на 1. годину, 2011/12. школске године поново уписан у 2. годину, самофинансирање, студијски програм Информациони системи и менаџмент, Информациони системи, током студија положио је испите из следећих предмета:

Р.бр.	Шифра	Назив предмета	Оцена	ЕСПБ	Фонд часова**	Датум
1.	Д00068	Методологија научно – истраживачког рада	10 (десет)	10	I:(45+0+60)	08.09.2010.
2.	Д00011	Квалитет софтвера – одабрана поглавља	10 (десет)	10	II:(45+0+60)	08.10.2010.
3.	Д00012	Интеракција човека и рачунара - одабрана поглавља	10 (десет)	10	II:(45+0+60)	08.10.2010.
4.	Д00001	Управљање подацима	10 (десет)	10	I:(45+0+60)	14.10.2010.
5.	Д00002	Развој информационих система	10 (десет)	10	I:(45+0+60)	20.11.2010.
6.	Д00033	Пословна интелигенција – изабрана поглавља	10 (десет)	10	III:(45+0+60)	10.03.2011.
7.	Д00021	Менаџмент информациони системи	10 (десет)	10	III:(45+0+60)	10.03.2011.
8.	Д00034	Мултимедијалне комуникације	9 (девет)	10	III:(45+0+60)	07.06.2011.
9.	Д00007	Моделовање предузећа	9 (девет)	10	III:(45+0+60)	10.08.2011.

* - еквивалентира/признат испит.

** - фонд часова је у формату (предавања+вежбе+остало).

Укупно остварено 90 ЕСПБ.

Општи успех: 9,78 (девет и 78/100) , по годинама студија (10,00, 9,50, /).

Студент није завршио факултет

Виши стручни сарадник за докторске студије

др Марина Јовановић-Миленковић

Поводом предлога Већа студијског програма докторских академских студија, а на основу члана 56. Статута Факултета, на седници Наставно-научног већа одржаној 11.07.2012. године, једногласно је донета следећа

ОДЛУКА

Усваја се Извештај Комисије за оцену научне заснованости пријављене докторске дисертације кандидата на докторским студијама Драгољуба Крнете под насловом »АУТОМАТИЗАЦИЈА ФИЗИЧКОГ ПРОЈЕКТОВАЊА СКЛАДИШТА ПОДАТАКА ПРОШИРЕНОГ *DATA VAULT* ПРИСТУПА«. За ментора се предлаже др Зоран Марјановић, ред. проф. ФОН-а.

Извештај се доставља Већу научних области техничких наука Универзитета у Београду.

Председник Наставно научног већа

Проф. др Милан Мартић

Доставити:

- Стручном сараднику за ПДС
- Ментору
- Вишем референту органа управљања и стручних органа
- Архиви

ПОДАЦИ О МЕНТОРУ

за кандидата Драгољуба Крнету

Име и презиме ментора: др Зоран Марјановић

Звање: ред. проф. ФОН-а

Списак радова који квалификују ментора за вођење докторске дисертације:

1. М. Вујасиновић, Н. Ивезић, Е. Barkmeyer, З. Марјановић: „**Semantic B2B-integration based on an Ontological Message Metamodel**“. Concurrent Engineering: Research & Applications, ISSN 1063-293X, IF 2009 0.959, Vol. 18. No. 3. Sept. 2010. pp. 219-232
2. М. Вујасиновић, Е. Barkmeyer, Н. Ивезић, З. Марјановић: „**Interoperable Supply-Chain Applications: Message Metamodel-based Semantic Reconciliation of B2B Messages**“. International Journal of Cooperative Information Systems, ISSN 0218-8430, IF 2008 0.714, Vol. 19. No. 1 & 2, Mar/June 2010. pp. 1-39
3. М. Вујасиновић, Н. Ивезић, В. Kulvatunyou, Е. Barkmeyer, М. Missikoff, F. Taglino, З. Марјановић, И. Милетић: „**Semantic Mediation for Standard-Based B2B Interoperability**“. IEEE Internet Computing, ISSN 1089-7801, IF 2008 2.309, Vol. 14, No. 1, 2010, Jan/Feb. 2010. pp. 52-63
4. М. Вујасиновић, Н. Ивезић, В. Kulvatunyou, Е. Barkmeyer, М. Missikoff, F. Taglino, З. Марјановић, И. Милетић: „**A Semantic-mediation Architecture for Interoperable Supply-chain applications**“. International Journal of Computer Integrated Manufacturing, ISSN 0951-192X, IF 2008 0.722, Vol. 22, No 6, Dec. 2008, pp. 549-561
5. М. Вујасиновић, З. Марјановић: „**Data Level Enterprise Applications Integration**“, LNCS 3812 (Lecture Notes in Computer Science), Special issue on “Business Process Management Workshops”, Volume Editors: Christoph Bussler and Armin Haller, Springer, 2006. pp. 390-396 ISSN 0302-9743, IF 2004 0.513. Напомена: Рад је штампан у електронском издању часописа Lecture Notes in Computer Science, на адреси http://dx.doi.org/10.1007/11678564_35.

Заокружити одговарајућу опцију(А,Б,В или Г):

А) У случају менторства дисертације на докторским студијама у групацији техничко-технолошких, природно-математичких и медицинских наука ментор треба да има најмање три рада са SCI, SSCI, AHCI или SCIE листе.

Б) У случају менторства дисертације на докторским студијама у групацији друштвено-хуманистичких наука ментор треба да има најмање три рада са релевантне листе научних часописа (Релевантна листа научних часописа обухвата SCI, SSCI, AHCI и SCIE листе, листу часописа које је Министарство за науку класификовало као M24 и додатну листу часописа коју ће, на предлог универзитета, донети Национални савет за високо образовање. Посебно се вреднују и монографије које Министарство науке класификује као M11, M12, M13, M14, M41 и M51.)

В) У случају израде докторске дисертације према ранијим прописима за кандидате који су стекли академски назив магистра наука ментор треба да има

пет радова (рефернци) које га, по оцени Већа научних области, квалификују за ментора односне дисертације.

Г) У случају да у ужој научној области нема квалификованих наставника, приложити одлуку Већа докторских студија о именовању редовног професора за ментора.

ДЕКАН ФАКУЛТЕТА

Датум 11.07.2012.

М.П.



NASTAVNO-NAUČNOM VEĆU FAKULTETA

Predmet: Izveštaj o podobnosti kandidata i naučnoj
zasnovanosti prijavljene doktorske disertacije

Odlukom nastavno-naučnog veća FON-a br. 3/64-19 od 30.05.2012. godine imenovani smo u Komisiju za pregled i odbranu pristupnog rada i ocenu naučne zasnovanosti prijavljene doktorske disertacije kandidata Dragoljuba Krnete, pod naslovom:

Automatizacija fizičkog projektovanja skladišta podataka proširenog Data Vault pristupa

i na osnovu toga podnosimo sledeći

I Z V E Š T A J

1. Podaci o kandidatu

Opšti biografski podaci

Dragoljub Krneta je rođen 06.01.1966. u selu Zebe opština Bosansko Grahovo, Bosna i Hercegovina. Osnovnu i srednju školu završio je u Bosanskom Grahovu kao nosilac diplome Ognjen Prica (pandan Vukove diplome u Srbiji). Završio je Elektrotehnički fakultet prvi stepen, smjer računari i automatika, na Univerzitetu u Banja Luci 1994. godine. Tehnički fakultet „Mihajlo Pupin“ u Zrenjaninu na Univerzitetu u Novom Sadu završava 2006. godine, a na istom fakultetu 2008. godine završava postdiplomske master studije sa prosečnom ocenom 9,23 i stiče zvanje Diplomirani inženjer poslovne informatike – master. Master rad u oblasti poslovne inteligencije pod mentorstvom prof. dr Dragice Radosav odbranio je sa ocenom 10.

Programiranjem, razvojem softvera i bazama podataka bavi se od 1989. Do 2002. godine autor je dvadesetak različitih softverskih poslovnih aplikacija koje su implementirane u više od 50 preduzeća. Od 2002. godine prelazi u kompaniju Lanaco Informacione tehnologije u Banja Luci (danas najveću IT kompaniju u BiH) na poziciju Menadžera sektora za razvoj softvera. Za vrijeme desetogodišnjeg rada u Lanacu, učestvovao je u svim značajnijim Lanaco softverskim projektima bilo kao projekt menadžer, projektant informacionog sistema i baze podataka, sistem analitičar ili programer. Pored upravljanja sektorom za razvoj softvera, zadužen je za obezbeđenje kvaliteta softverskih proizvoda prema aktuelnim standardima, pružanje smjernica u procesu razvoja softvera, praćenje procesa održavanja softvera, ugovaranje prodaje aplikativnog softvera, te saradnju sa drugim sektorima u realizaciji kompleksnih informacionih sistema. Posjeduje razne sertifikate u informacionim tehnologijama od kojih su najznačajniji: Microsoft Certified

Professional (MCP) i Microsoft Certified Database Administrator (MCDBA). Učestvovao je na Microsoft Worldwide Partner Conference u Denveru 2007. i u Washingtonu 2010. godine. Član je asocijacije Institute of Electrical and Electronics Engineers (IEEE).

Kandidat Dragoljub Krneta je na doktorskim studijama Fakulteta organizacionih nauka položio svih devet ispita (sa prosečnom ocenom 9,67) predviđenih planom i programom doktorskih studija na smeru Informacioni sistemi i ostvario ukupno 90 ESPB bodova.

Rb	Predmet	Ocena	ESPB
1	Upravljanje podacima	10	10
2	Razvoj informacionih sistema	10	10
3	Interakcija čoveka i računara	10	10
4	Kvalitet softvera	9	10
5	Metodologija naučno-istraživačkog rada	10	10
6	Modelovanje preduzeća	9	10
7	Menadžment informacioni sistemi	10	10
8	Multimedijalne komunikacije	9	10
9	Poslovna inteligencija	10	10

90

Spisak objavljenih radova

A. Master rad

Dragoljub Krneta: *Poslovna primena Microsoft alata i tehnologija*, Oblast: *Poslovna informatika*, Mentor: *Prof. dr Dragica Radosav*, Univerzitet u Novom Sadu, Tehnički Fakultet „Mihajlo Pupin“, Zrenjanin, juli, 2008

B. Spisak naučnih i stručnih radova

Kandidat Dragoljub Krneta je autor ili koautor više naučnih i stručnih radova iz oblasti razvoja softvera, baza podataka, poslovne inteligencije i Data Warehouse sistema.

1) Radovi objavljeni na domaćoj konferenciji:

- Radulović B., Krneta D., *Modeli za ocenu kvaliteta softvera*, SYMOPIS 2007., Zlatibor, ISBN: 978/86/7680/124/4, <http://www.symopis.fon.bg.ac.yu>

2) Radovi objavljeni na međunarodnoj konferenciji:

- Krmeta D., *Informacioni sistem Fonda za zdravstveno osiguranje Republike Srpske*, SEFICT (South East Europe Forum ICT) 2006., Beograd, <http://www.sefict.org>
- Krmeta D., *Održavanje softvera*, SEFICT (South East Europe Forum ICT) 2007., Herceg Novi, <http://www.sefict.org>
- Krmeta D., Radulović B., Radosav D., *Business Intelligence u zdravstvenom informacionom sistemu*, Infoteh 2008., Jahorina, ISBN: 99938-624-2-8, <http://www.infoteh.rs.ba>
- Krmeta D., Radosav D., Radulović B., *Realization Business Intelligence in Commerce using Microsoft Business Intelligence*, SISY 2008., Subotica, ISBN: 978-1-4244-2407-8, IEEE Catalog Number: CFP0884C-CDR, <http://bmf.hu/conferences/sisy2008>
- Krmeta D., Radosav D., *Realization Business Intelligence using Microsoft SQL Server 2008*, Conference on Information and Communication Technologies 2008., Mostar
- Krmeta D., Radosav D., *HL7 Protocol*, Infoteh 2009., Jahorina, ISBN: 99938-624-2-8, <http://www.infoteh.rs.ba>
- Banović N., Rekić Z., Krmeta D., *Značaj i uloga elektronskog školskog dnevnika u vaspitno-obrazovnom procesu*, ITEO konferencija, 2009, Banja Luka
- Krmeta D., Opačić V., Bobanac Z., Banović N., *Sistem poslovne inteligencije u Fondu zdravstvenog osiguranja*, Infoteh 2010., Jahorina, ISBN: 99938-624-2-8, <http://www.infoteh.rs.ba>
- Krmeta D., Marjanović Z., *Mogućnost integracije metodoloških pristupa razvoja informacionog sistema i implementacije gotovog ERP rješenja*, Infoteh 2011., Jahorina, ISBN: 99938-624-2-8, <http://www.infoteh.rs.ba>

3) Softveri:

Kandidat Dragoljub Krmeta je učestvovao u više desetina softverskih projekata kao projekt menadžer, projektant informacionog sistema i baze podataka, sistem analitičar ili programer. U sledećoj tabeli navodimo najznačajnije softverske projekte.

Projekat i naručilac	Vrijeme trajanja Pozicija na projektu	Kratak opis poslova koje je obavljao
Poslovni informacioni sistem Fond za zdravstveno osiguranje Republike Srpske	2006-2012 Projekt menadžer za održavanje softvera	Upravljanje održavanjem softvera; planiranje i analiza održavanja softvera; izrada prateće dokumentacije u procesu održavanje softvera.
Poslovni informacioni	2010	Analiza sistema; projektovanje

sistem Radio-televizija Republike Srpske	Projektant informacionog sistema	baze podataka; razvoj standardne forme; izrada projektne dokumentacije
Sistem poslovne inteligencije Fond za zdravstveno osiguranje Republike Srpske	2009 Projekt menadžer	Upravljanje projektom; analiza postojećeg sistema; razvoj Data Warehouse i BI sistema; razvijanje standarde za korištenje aplikacije; izrada projektne dokumentacije.
Informacioni sistem za obračun, naplatu i tarifiranje isporučene toplotne energije Toplana AD Banja Luka	2008-2009 Projekt menadžer	Upravljanje projektom; Učešće u razvoju softvera, analiza postojećeg sistema; projektovanje baze podataka; razvoj standardne forme.
Informacioni sistem za vođenje osnovnih sredstava Vlada Republike Srpske	2009 Projektant informacionog sistema	Analiza sistema za vođenje osnovnih sredstava; projektovanje baze podataka; razvoj aplikacije.
Softver za obračun zarada Rudnik i Termoleketrana Gacko	2008 Projektant IS	Analiza sistema za vođenje obračun zarada; projektovanje baze podataka; razvoj aplikacije.
Poslovni informacioni sistem Elektroprenos BiH	2007 Projekt menadžer	Upravljanje projektom; učešće u razvoju softvera, analiza postojećeg sistema; razvoj baze podataka; upravljanje i koordinacija u procesu razvoja softvera.
Poslovni informacioni sistem preduzeća (moduli: nabavke, finansijsko, materijalno, osnovna sredstva, ljudski resursi, zarade, proizvodnja) Lanaco Banja Luka (sistem implementiran u više od 100 preduzeća)	2004-2010 Projekt Menadžer, Sistem analitičar, DB dizajner	Upravljanje projektom; učešće u razvoju softvera, analiza sistema; razvoj baze podataka; razvoj standardne forme; odgovornost za upravljanje i koordinaciju u procesu razvoja softvera; odgovornost za implementaciju sistema.
Poslovni informacioni sistem Integral Inženjering Corporation Laktaši	2005-2006 Projekt Menadžer, Sistem analitičar, DB dizajner,	Upravljanje projektom; učešće u razvoju softvera, analiza postojećeg sistema; razvoj baze podataka; odgovornost za upravljanje i koordinaciju u

	Programer	procesu razvoja softvera.
SOTAC (SOcial Sector Technical Assistance Project) Vlada Republike Srpske Vlada Federacije BiH (sistem implementiran u 164 opštine u BiH)	2003-2004 Projektant informacionog sistema	Analiza sistema; projektovanje baze podataka; razvoj aplikacije; implementacija sistema i obuka korisnika
Poslovni informacioni sistem Fond za zdravstveno osiguranje Republike Srpske (sistem implementiran u 62 opštine u Republici Srpskoj)	2002-2004 Projekt menadžer Projekt Menadžer, Sistem analitičar, DB dizajner, Programer	Upravljanje projektom; učešće u razvoju softvera, analiza sistema; razvoj baze podataka; razvoj standardne forme; odgovornost za upravljanje i koordinaciju u procesu razvoja softvera.

Ocena podobnosti kandidata za rad na predloženoj temi

Kandidat Dragoljub Krneta se već dugi niz godina bavi programiranjem, razvojem softvera i bazama podataka. Učestvovao je u više desetina softverskih projekata. Uža oblast interesovanja kandidata su baze podataka, skladišta podataka i sistemi poslovne inteligencije. Na doktorskim studijama Fakulteta organizacionih nauka položio svih devet ispita predviđenih planom i programom studija i ostavrio 90 ESPB bodova.

Kao što se u priloženoj biografiji može videti, zadnjih deset godina je uključen u naučno-istraživački rad kroz učešće na IT konferencijama i u okviru IT projekata na kojima je učestvovao. Učestvovao je u dosta značajnih projekata koje je u okviru IT kompanije Lanaco Informacione tehnologije radio za preduzeća, državne organe i institucije, stekavši izuzetno iskustvo u analizi, projektovanju i implementaciji složenih informacionih sistema. O tome govori i veliki broj različitih softverskih rešenja nabrojanih u referencama.

Na osnovu svega navedenog, smatramo da kandidat Dragoljub Krneta poseduje potrebno znanje i iskustvo za izradu doktorske disertacije na temu „**Automatizacija fizičkog projektovanja skladišta podataka proširenog Data Vault pristupa**“.

2. Predmet i cilj istraživanja

Predmet istraživanja ovog rada u širem smislu obuhvata sledeće oblasti: baze podataka, skladišta podataka, poslovnu inteligenciju i podršku odlučivanju. U užem smislu, predmet istraživanja je proces dizajniranja, odnosno izrade modela skladišta

podataka. Normalizovane baze podataka su projektovanje za često dodavanje, brisanje i izmjenu podataka, dok su dimenzione baze projektovanje za česte upite nad podacima i minimalno ažuriranje podataka. Kod dimenzionih baza podataka relacije su podjeljene u dvije kategorije: tabelu mjera ili činjenica (*eng. fact table*) i više tabela dimenzija (*eng. dimension tables*). Tabela mjera ima dvije vrste atributa: ključne attribute koji su ujedno spoljni ključevi pomoću kojih je tabela činjenica povezana sa tabelama dimenzija i same činjenice koje sadrže podatke koji se analiziraju. Tabele su najčešće spojene u šemu zvijezde (*eng. star schema*).

Data vault je koncept skladišta podataka čiji je autor D.Linstedt. Data vault se realizuje hibridnim pristupom između podataka u trećoj normalnoj formi i zvjezdaste šeme, te omogućuje praćenje i rekonstrukciju historijskih podataka. D.Linstedt uvodi *hub*, *link* i *satelit* entitete. Hub predstavlja primarni ključ, link entiteti omogućuje transakcije između hubova-a, dok satelitski entiteti daju kontekst hub primarnom ključu.

Centralni problem istraživanja se odnosi na mogućnost automatizacije projektovanja skladišta podataka na osnovu fizičkog modela transakcione relacione baze podataka. Drugim riječima, ispitaće se mogućnost automatskog kreiranja tabela mjera i dimenzija, odnosno hub, link i satelit entiteta u data vault konceptu.

Osnovni cilj pristupnog rada je da se sagleda problem neautomatizovane izrade dizajna skladišta podataka na osnovu šire literature, te istraži mogućnost automatizacije projektovanja skladišta podataka baziranog na Data Vault konceptu uz minimalnu interakciju sa korisnikom.

Ostali ciljevi doktorske disertacije su:

- ispitati mogućnost automatizacije ETL procesa u dijelu prikupljanja, obrade i učitavanja internih podataka iz transakcione baze,
- istražiti mogućnost uticaja automatizacije dizajna skladišta podataka na brzinu i kvalitet dobijanja informacija o poslovnim procesima i potrebama korisnika za informacijama,
- ispitati mogućnost smanjenja troškova i vremena implementacije skladišta podataka,
- demonstrirati upotrebnu vrednost automatizacije za profesionalnu praksu u domenu zdravstva.

Pregled dosadašnjih rezultata istraživanja u domenu na koji se disertacija odnosi

Skladište podataka (*eng. Data Warehouse, DW*) predstavlja podlogu za implementaciju sistema poslovne inteligencije (*eng. Business Intelligence, BI*), integraciju podataka i podršku poslovnom odlučivanju. Poslovna inteligencija je generički termin za opisivanje uticaja unutrašnjih i vanjskih informacija organizacije za donošenje boljih poslovnih odluka [Kibmall, 2002]. Skladište je predmetno

orijentisana, integrisana, stalna i vremenski promjenjiva kolekcija podataka za podršku menadžmentu kod odlučivanja [Inmon, 1992]. Kao specifična vrsta baza podataka, skladište podataka su se razvila krajem 80-tih i početkom 90-tih godina prošlog vijeka. Ralph Kimball je 1986. godine napravio tzv. Red Brick Systems koji je bio zasnovan na relacionoj bazi podataka optimizovan za skladištenje podataka koji je u to vrijeme imao performanse i do deset puta bolje u odnosu na tadašnje baze podataka. Eftimie i Briquez su 1984. godine su predstavili tehnike modelovanja i dizajna skladišta informacija (*eng.* Information Warehouse) radi podrške menadžmentu kompanija u donošenju odluka. Na projektu integracije informacionog sistema brodogradilišta 3.maj u Rijeci 1986. god. i za potrebe informacionog sistema Privredne komore Jugoslavije 1988. godine, tim pod rukovodstvom V. Jovanovića i Z. Marjanovića je nezavisno razvio sistem za integraciju podataka koristeći multidimenzionalan dizajn. Devlin i Murphy su 1988. godine objavili prvi rad koji opisuje arhitektura skladišta podataka. Bill Inmon je 1992. godine objavio prvu knjigu o skladištenju podataka. Osnovni razlog koji je doprineo nastanku i razvoju skladišta podataka bila je nemogućnost tadašnjih baza podataka (*eng.* Database, DB), OLTP informacionih sistema da podrže sve veće zahtjeve za obradu i analizu poslovnih informacija uključujući istorijske i eksterne, te zahtjeve za estimacije i predviđanja. Prema podacima istraživačke agencije Gartner iz kraja 2010. godine, sistemi poslovne inteligencije se nalaze na prvom mjestu od prvih deset tehnoloških prioriteta kompanija u SAD-u. Od 2000. godine razvoj i investicije u DW prevazišle su one u DB sistemima i taj se trend ubrzano nastavlja sa eksplozivnim razvojem storage sistema (prvo SAN, a sada Virtualization i Cloud).

Kod transakciono orijentisanih relacionih baza podataka (OLTP – On Line Transaction Procesing) izrađuje se najčešće normalizovan model podataka, što povlači složen način pristupa podacima. Za skladišta podataka (posebno na strani korisnika) uglavnom se izgrađuje dimenzioni model kojim se dobija bolja mogućnost vizualizacije podataka. Najčešće se koristi tzv. zvjezdasta (*eng.* star) šema podataka. Postupak prikupljanja i obrade podataka za potrebe punjenja skladišta podataka naziva se ETL (Extraction, Transformation and Loading) proces. Data mart je oblik skladišta podataka koji najčešće sadrži već sumirane podatke na određenim organizacionim dijelovima kompanije ili po određenoj grupi poslovnih procesa. Data mart je mjesto gdje krajnji korisnik ima direktan pristup i kontrolu njegovih analitičkih podataka. Finansije imaju svoje podatke, marketing i prodaja svoj data mart, itd. Svaki data mart obično drži znatno manje podataka nego skladište podataka [Inmon, 2002]. Kod arhitekture skladišta podataka danas postoje dva dominantna pristupa. Jedan pristup, poznat je pod nazivom CIF (Corporate Information Factory) arhitektura, zastupa Bill Inmon koji smatra da skladište podataka treba da bude normalizovano i da sadrži atomske podatke u trećoj normalnoj formi, te da koristi višedimenzionalne data mart-ove. Drugi pristup, poznat pod nazivom BUS arhitektura, zastupa Ralph Kimbal koji tvrdi da i samo skladište treba da ima višedimenzionalnu strukturu.

Data Vault koncept skladišta podataka koji je osmislio Daniel Linstedt objavivši sredinom 2002. godine prvi od pet članaka pod nazivom Data Vault Series. Ostali članci su objavljeni do početka 2005. na sajtu <http://www.tdan.com/view-articles/5054/>. Data Vault se pojavio kao odgovor na probleme koji su pratili razvoj skladišta podataka dizajniranih u trećoj normalnoj formi i star šemi, nastojeći iskoristiti najbolje elemente jednog i drugog koncepta. Performanse i druge slabosti 3NF i zvijezda shema (kada se koristi za skladište podataka) počele su se pojavljivati u 90-im kada se količina podataka počela povećavati. Data Vault je projektovan za prevazilaženje tih nedostataka zadržavajući prednosti 3NF i zvijezda shema arhitekture [Linstedt, 2002]. Data Vault predstavlja detaljno orijentisani, istorijski praćen i jedinstveno povezani skup normalizovanih tabela koji podržavaju jedan ili više funkcionalnih područja poslovanja. To je hibridni pristup koji obuhvata najbolje osobine između 3. normalne forme (3NF) i zvijezda šeme. Dizajn je fleksibilan, skalabilan, dosljedan i prilagodljiv potrebama poduzeća. To je model podataka koji je projektovan posebno za potrebe skladišta podataka organizacije [Linstedt, 2002-2005]. Linstedt uvodi *hub*, *link* i *satelit* entitete. Hub predstavlja poslovni ključ, link entiteti omogućuje transakcije između hubova, dok satelitski entiteti daju kontekst hub primarnom ključu.

Data Vault koncept naglašava potrebu da se ostavi trag odakle i kad su podaci došli u bazu podataka [Linstedt, 2010]. Osim toga, ovaj koncept je osmišljen da se model podataka može lako promjeniti u skladu sa promjenama u poslovnom okruženju [Damhof, 2008], te kako bi se omogućilo paralelno i brzo učitavanje podataka koliko je god to moguće [Hammergren, Simon, 2008]. Dizajn Data Vault sistema je jednostavan sa minimalnim brojem komponenti kojeg čine Hub, Link i Satelit entiteti. Kod Data Vault sistema u praksi se primjenjuje i minimalni broj pomoćnih tabela tzv. reference data.

Skladišta podataka se rade različitim metodološkim pristupima, uzimajući u obzir specifičnosti svake organizacije. Metodologija izgradnje skladišta podataka (uključujući popularne kao Kimbell DWLC Toolkit, sa varijantom prilagođenom za Microsoft SQL Server 2008, ali ni najnovije kao [Corr, 2011]) nije u potpunosti formalizovana, već sadrži niz preporuka kojih bi se trebalo držati. Izgradnja skladišta podataka uglavnom počinje definisanjem korisničkih zahtjeva i upoznavanjem projektanata sistema sa njihovim željama i potrebama. Nakon analize zahtjeva korisnika i upoznavanja sa strukturom transakcione baze podataka, prelazi se na izradu logičkog modela, a zatim i fizičkog modela podataka za DW.

Međutim, jedan problem je što proces dizajniranja skladišta podataka i izrade dimenzionog modela nije automatizovan. Dimenzioni model sa tabelama mjera i dimenzija se u većini slučajeva radi ručno na osnovu analize zahtjeva korisnika i analize transakcionog sistema, bilo da se radi o Inmon-ovoj ili Kimbal-ovoj metodologiji ili se radi o primenama po Data Vault konceptu. Problem neautomatizovane izrade dimenzionog modela skladišta podataka daje motivaciju da se istraži mogućnost automatizacije dizajna skladišta podataka, odnosno

automatskog generisanja i prepoznavanja mjera i dimenzija na osnovu dizajna i fizičkog modela transakcione baze podataka uz minimalnu interakciju sa korisnikom. Osim toga, ovakvo istraživanje otvara perspektive za istraživanje automatizacije ETL procesa u dijelu prikupljanja, obrade i učitavanja internih podataka iz transakcione baze, te posebno podršci fizičkog projektovanja DW/DM, pitanju particija po vremenu i/ili lokaciji kao i distribuciji kopija i materijalizovanim agregacijama za data martove.

[Golfarelli et al, 1998.] predlaže konceptualni model skladišta podataka i poluautomatski sekvencijalni hibridni pristup dizajna skladišta podataka koji proizlazi iz E/R dokumentacije. Dimensional Fact Model (DFM) prevazilazi jaz između činjenica i šeme skladišta podataka. Logički dizajn se treba zasnivati na procjeni očekivanog opterećenja i količine podataka. U radu se navodi mogućnost kombinovanjem posebnih dimenzija u jednu definisanjem kompozitnih ključeva i formiranje degenerisanih dimenzija kojih nema u E/R dijagramu.

[Phipps i Davis, 2002.] predlažu i ilustruju algoritam za automatsko generisanje konceptualne šeme koristeći šemu operativne baze podataka. Kao polazište se uzima data-driven pristup, a kasnije se prelazi na demand-driven (sekvencijalni hibridni pristup). Polazna konceptualna šema je stvorena korištenjem Multidimensional E/R modela. Algoritam koristi numerička polja i odnosa između entiteta kao temelj za stvaranje Multidimensional E/R šeme. Algoritam se primjenjuje za svaki model podataka gdje se podaci mogu se razdijeliti u numeričke, datum/vrijeme i tekstualne tipove podataka. Osim toga, ovaj pristup se može koristiti za procjenu kandidata za konceptualne šeme pomoću korisničkih upita. Rad predstavlja značajan napredak u automatizaciji dizajna skladišta podataka.

[Peralta et al, 2003] predstavlja korak naprijed ka automatizaciji skladišta podataka baziran na pravilima primjenom postojećih znanja u dizajnu skladišta podataka. Predložena pravila se ugrađuju u dizajn strategiju koja se pokreće u skladu sa zahtjevima i izvodom podataka, te na osnovu šeme transakcione baze.

[Dori et al, 2005] prikazuje Object Process Methodology (OPM) kao metod za dobijanje smjernica za transformaciju transakcionog sistema u šemu skladišta podataka. Ovaj metod se koristi za identifikaciju relevantnih entiteta za poslovni proces čijim korištenjem dizajner može izabrati procese koji na najbolji način reprezentuju strukturu skladišta podataka. Pristup koji se predlaže u radu je poluautomatizovan hibridni pristup dizajniranja skladišta podataka jer u određenoj mjeri zahtjeva učešće korisnika.

[Romero i Abelo, 2007] predstavljaju novi pristup za automatizaciju dizajna skladišta podataka. U radu se predlaže poluautomatska metoda s ciljem kako bi pronašli poslovne višedimenzionalne koncepte iz domene ontologije koji predstavljaju

različite i potencijalno heterogene podataka. Uz korištenje ontologije i obrazloženje alate koje ontologije koristi u potrazi za višedimenzionalnim uzorcima, provedena je simulacija stvarnog slučaja za provjeru algoritam. Predstavljeno je i teoretsko proučavanje algoritama. Ovaj pristup je u stanju integrisati podatke iz heterogenih izvora podataka koji opisuju njihove domenu kroz ontologije. Jedan od najperspektivnijih područja gdje se metoda može primjeniti je semantički web, što doprinosi vađenje i integrisanju vanjskih podataka sa web-a u skladište podataka.

[Zepeda et al, 2010] predlaže polu-automatski proces za izgradnju skladišta podataka zasnovan na MDA. Proces se sastoji od četiri koraka. Prvi se bavi sa prikupljanjem i konsolidovanjem zahtjeva korisnika. Ostali koraci polaze od formalnih i strukturnih informacije iz šeme izvorne baze podataka i mogu biti podržani od strane neke automatizacije tehnike. Izbor MDA kao vodeće metodologije opravdava potrebu iskorištavanja šeme baze. U radu se koristi alat koji je razvijen za podrška MDA razvoju skladišta podataka i koji može bitno usmjeriti dizajnera prema efikasnom rješenju u skladu sa zahtjevima korisnika, te ETL proces učiniti jednostavnijim.

[Nazri et al, 2010] predlaže novu metodu za projektovanje skladišta podataka koja se zasniva na modelu izvora podataka. Predlaže se poluautomatski alat koji koristi leksičke ontologije kao bazu znanja. Nakon što identifikuju činjenice i dimenzije, sljedeći korak je generisanje višedimenzionalnog modela skladišta podataka sa minimalnim učešćem korisnika. Korištenje predložene metode je pokazana je na primjeru koristeći WordNet semantički leksikon kao znanje domena, odnosno kao pomoć pri identifikaciji potencijalnih mjera i dimenzija.

[Zekri et al, 2011] uvodi novi pristup za poluautomatsko oblikovanje multi-dimenzionalnog dijagrama. Conceptual Data Model (CDM) se najprije prevodi na multivarijantni patern, a zatim se rafinira pomoću zahtjeva korisnika. Oba koraka koriste grafove kao formalizam za predstavljanje modela podataka za donošenje odluka. Graf se sastoji od dvije vrste čvorova: pravougaonih čvorova kao potencijalnim činjenice i kružnim čvorovima kojima odgovaraju dimenzije. Izrada modela ovakvim pristupom može biti znatno automatizovati na osnovu modela izvornih podataka

U pregledanoj literaturi koja se odnosi na dizajn Data Vault sistema [Linstedt 2002-2010, Damhof 2008, Hamergen i Simon 2010, Graziano 2011, Casters, Bounan i VanDongen 2010] je pomenuta mogućnost određene automatizacije dizajna skladišta podataka. Međutim, u literaturi nije nađen konkretan postupak niti formalizacija automatizacije dizajna Data Vault sistema. Postupak automatizacije dizajna skladišta podataka nije nimalo jednostavan proces, jer pored identifikacije šeme izvorne baze podataka, potrebno je formalizovati zahtjeve krajnjih korisnika. Određeni koraci ipak moraju biti ručno izvedeni. Najčešće je to postupak identifikacije poslovnih ključeva i

poslovnih mjera. Identifikacija link i satelit tabela u Data Vault konceptu, odnosno izbor dimenzija se može u značajnoj mjeri automatizovati. Posjedovanje procedura i alata za što veću automatizaciju dizajna skladišta podataka u znatnoj mjeri smanjiće vrijeme razvoja skladišta podataka.

3. Polazne hipoteze

U radu se postavlja sledeća **generalna hipoteza** koja obuhvata preliminarno i teoretsko određenje predmeta istraživanja:

Hipoteza 1: Proces projektovanja skladišta podataka baziranog na Data Vault konceptu i izrade dimenzionog modela skladišta podataka može se u određenoj mjeri automatizovati na osnovu fizičkog modela transakcione baze podataka.

Iz generalne hipoteze izvode se sledeće **posebne hipoteze** koje obrađuju dijelove predmeta istraživanja:

Hipoteza 2: Primjenom automatizacije u dizajnu skladišta podataka, podiže se kvalitet DW sistema, te smanjuju vrijeme i troškovi implementacije skladišta podataka i povezanog sistema poslovne inteligencije.

Hipoteza 3: Automatizacija izrade dimenzionog modela skladišta podataka doprinosi automatizaciji ETL procesa.

Istraživanjem će se dokazati ili odbaciti postavljene hipoteze.

4. Naučne metode istraživanja

Osnovni metod istraživanja u pristupnom radu je sakupljanje i proučavanje dostupne literature, njena analiza i sistematizacija, a sve to s ciljem da se pokaže opravdanost i korisnost razvoja novog modela automatizacije projektovanja skladišta podataka.

U izradi rada od opštih naučnih metoda koristiće se metoda modelovanja, analitičko-deduktivna i statistička metoda. Metoda modelovanja će se koristiti prilikom izrade modela skladišta podataka i praktične provjere rezultata istraživanja u oblasti zdravstva. Kod analize podataka o postojećim rješenjima u oblasti skladišta podataka će se koristiti analitičko-deduktivna metoda. Analiza dobijenih rezultata će se dijelom obaviti preko standardnih statističkih metoda. Pored navedenih opštih metoda, koristiće se i posebni metodološki pristupi kao što su generalizacija i specijalizacija vezani za primjene u oblastima baza podataka na koje se odnosi istraživanje.

5. Očekivani naučni doprinos

Osnovni doprinos pristupnog rada ogleda se u definisanju sistematskog pregleda proučavane oblasti kao i **predlaganje originalnog pristupa za automatizaciju projektovanja skladišta podataka baziranog na Data Vault konceptu**. Razvijeni pristup će moći da se koristi kao sredstvo za evaluaciju efikasnosti i moguće ga je primeniti na različite tipove organizacija i različite sisteme za upravljanje bazama podataka.

Očekivani naučni **doprinos doktorske disertacije** će biti u unapređenju procesa dizajniranja skladišta podataka, odnosno formalnom opisu modela i metoda koji omogućuju što veću automatizaciju izrade modela skladišta podataka. Osim toga, **rad će dati celokupan integrisan pristup automatizaciji projektovanja skladišta podataka**. Rezultati istraživanja će doprinjeti podizanju kvaliteta DW sistema, te smanjenju troškova i vremena implementacije DW sistema i lakšem otkrivanju znanja o poslovnim procesima i potreba korisnika za informacijama. Profesionalni doprinos od posebnog interesa će predstavljati upotrebna vrijednost automatizacije dizajna skladišta podataka u oblasti zdravstva.

6. Okvirna plan istraživanja i struktura disertacije po poglavljima

U sledećoj tabeli je dat okvirni plan istraživanja:

Rb	Aktivnost	Vremenski period
1.	Analiza i opis osnovnih aspekata skladišta podataka	juli, avgust 2012.
2.	Istorijski razvoj skladišta podataka	
3.	Analiza ETL procesa	
4.	Analiza relacionog i multidimenzionog dizajna	
5.	Skladišta podataka u sistemu poslovne inteligencije	
6.	Analiza arhitekture skladišta podataka	
7.	Metodologije razvoja skladišta podataka	
8.	Prostorna i vremenska skladišta podataka	
9.	Značaj i uloga Data vault koncepta	septembar, oktobar 2012.
10.	Najznačajniji dobavljači u oblasti skladišta podataka na svetskom tržištu	
11.	Pregled najnovijih istraživanja u oblasti automatizacije projektovanja skladišta podataka u Data Vault pristupu	
12.	Kritički osvrt na dosadašnje rezultate istraživanja	
13.	Identifikacija i formulacija problema	novembar, decembar 2012.; januar 2013.
14.	Istraživanje i prijedlog rješenja	
15.	Demonstracija upotrebne vrijednosti rezultata istraživanja na praktičnom primjeru u oblasti zdravstva	februar, mart, april 2013.
16.	Zaključna razmatranja	maj 2013.

Okvirna struktura doktorske disertacije

U uvodnom poglavlju biće prezentovani identifikovani problemi, kao i predmet i cilj istraživanja, te hipoteze i metode istraživanja.

U drugom dijelu pristupnog rada razmatraju se osnovni aspekti skladišta podataka. U ovom dijelu se daje nekoliko definicija skladišta podataka, sadržaj i ciljevi izgradnje skladišta podataka, te odnos poslovne inteligencije i skladišta podataka i njihov značaj za podršku odlučivanju. Dalje se razmatraju osnovni elementi multidimenzionalnog dizajna, ETL procesa i načina prezentacija podataka korisniku informacija. U drugom dijelu je dat pregled i karakteristike nekoliko arhitektura skladišta podataka i osnovni aspekti Data Vault koncepta skladištenja podataka.

Najnovija istraživanja u oblasti dizajna skladišta podataka razmatraju se u trećem dijelu pristupnog rada. Za pronalaženje literature, u ovom dijelu su korišteni najpoznatiji pretraživači naučnih radova sa konferencija i časopisa kao i sugestije projekatana i istraživača koji duže rade u oblasti DW radi kompletiranja slike i ekspertne ocjene relevantnosti smjernica prije detaljnog čitanja i ličnog istraživanja. U ovom dijelu rada se daje kritički osvrt na dosadašnje rezultate istraživanja, identifikuje se i analizira problem automatizacije dizajna skladišta podataka, te se daje prijedlog istraživanja da bi se došlo do rješenja.

U četvrtom dijelu se ispituje mogućnost automatizacije projektovanja skladišta podataka uz korištenje (u kontekstu industrijski veoma relevantnog) konkretnog sistema za upravljanje bazama podataka Microsoft SQL Server 2008 R2. Postoji nekoliko razloga za odabir ovog sistema kao što su: brojnost implementacija, raspoloživost uključujući cijenu, te intenzivan razvoj (sada već verzija 2012). Tehnički odlična alternativa Oracle bila bi od većeg značaja samo za najveće i finansijski najsnažnije korisnike.

U petom dijelu se pokazuje praktična provjera rezultata istraživanja, te razmatra upotrebnost automatizacije dizajna skladišta podataka u oblasti zdravstva i zdravstvenog osiguranja.

Zaključak, naučni doprinosi i pravci daljih istraživanja biće dati u šestom poglavlju. Na kraju rada se daju literatura i eventualno prilozi.

7. Zaključak i predlog

Iz izloženog se može zaključiti da kandidat Dragoljub Krneta ispunjava sve uslove predviđene Zakonom o visokom obrazovanju za odobrenje izrade doktorske disertacije pod naslovom „**Automatizacija fizičkog projektovanja skladišta podataka proširenog Data Vault pristupa**“.

Dragoljub Krneta poseduje neophodna znanja iz oblasti projektovanja složenih informacionih sistema i modelovanja poslovnih procesa, ogromno iskustvo stečeno aktivnim učešćem na više desetina razvojnih projekata iz oblasti informacionih sistema, kao i značajan istraživački kapacitet.

Tema pripada užoj naučnoj oblasti Informacioni sistemi, vrlo je aktuelna i na istim i sličnim problemima radi značajan broj istraživača širom sveta.

Na osnovu navedenog, komisija predlaže Nastavno-naučnom veću da prihvati predloženu temu i odobri izradu prijavljene doktorske disertacije.

Za mentora predlažemo Prof dr Zorana Marjanovića, redovnog profesora na Fakultetu organizacionih nauka Univerziteta u Beogradu.

Komisija

dr Zoran Marjanović,
redovni profesor FON-a, Univerzitet u Beogradu

dr Milija Suknović,
redovni profesor FON-a, Univerzitet u Beogradu

dr Vladan Jovanović,
redovni profesor na Georgia Southern University,
SAD