

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Подобност теме и кандидата **Гуме Абдулкадер Лакшена**, дипл.инж. (*Guma Abdulkhader Muntaser Lakshen, M.Sc.*) за израду докторске дисертације

Одлуком бр. **5054/13-1** од **23.05.2019.** године, именовани смо за чланове Комисије за оцену подобности теме и кандидата **Гуме Абдулкадер Лакшена** за израду докторске дисертације и научне заснованости теме „**Окружење за анализу и оцену квалитета великих и повезаних података**“ односно “*A framework for analysis and quality assessment of big and linked data*”.

На основу материјала приложеног уз Захтев кандидата, Комисија подноси следећи

ИЗВЕШТАЈ

1. Подаци о кандидату

1.1. Биографски подаци

Гума Абдулкадер Лакшен рођен је 1. јануара 1963. у месту Ал-Асаб, код Триполија у Либији. Основно и средњу школу је завршио од 1969 до 1982. У 1983. години Гума Абдулкадер Лакшен је добио стипендију за студије у Енглеској од стране новоосноване либијске компаније ЛИСЦО за гвожђе и челик. Средњу школу завршава у Енглеској 1985. године са одличним резултатима посебно у математици. Студије продужава на одсеку за компјутерски инжењеринг на колеџу у Кардифу, на универзитету у Велсу, 1985-1989. У периоду од 2002 до 2006, Гума Абдулкадер Лакшен студира и завршава мастер студије на смеру за компјутерске системе на Академији *Libyan Higher academy* са оценом 3.79/4.00.

Гума Абдулкадер Лакшен је своју каријеру започео радећи у ЛИСЦО-у од 1989. до 1997. године, током ових година углавном је радио као програмер, шеф компјутерске секције, и менаџер одељења за рачунарске системе (ИКТ Менаџер, *Libyan Iron and Steel Company, Misurata, Libya*). У периоду од 1998 до 2013. години радио је на следећим позицијама:

- ИКТ Менаџер (2000-2002) *Al-Maya medical factory, GMPMSCO, Al-Maya, Libya*
- Менаџер (2007-2009) *Al-rabta Pharmaceutical factory, GMPMSCO Libya, Al-rabta, Libya.*
- Менаџер логистике (2009-2011) *Libyan Pharmaceutical/medical Company, GMPMSCO, Tripoli, Libya.*
- Лидер одбора за међународне односе и инвестиције - *GMPMSCO*, где је сарађивао са компанијама из области медицине и фармације (*Hemofarm-Serbia, International PharmaChem-Italy, Oman Medical Supplies and Services Co. LLC- Oman*)

- Аналитичар података (2006-2011) *Almadena media and information center, Ministry of information and telecommunications, Tripoli, Libya.*
- Руководилац одељења / Професор (2011-2013). *Zintan teaching college, Al-Jabal Al-Gharbe University, Zentan, Libya.*

Аутор или коаутор је већег броја техничких решења насталих у оквиру комерцијалних пројекта, као и више радова публикованих у зборницима конференција националног, регионалног и међународног значаја.

Говори течно енглески језик.

1.2. Стечено научноистраживачко искуство

Гума Абдулкадер Лакшен (5054/2013) је уписао докторске студије, модул Софтверско инжењерство у 2013. години. Положио је све испите предвиђене планом и програмом Електротехничког факултета Универзитета у Београду са просечном оценом 9.5 и испунио све услове за пријаву докторске дисертације који су предвиђени Законом о високим образовању Р. Србије и Правилником о докторским студијама ЕТФ-а у Београду. Предмети на докторским студијама које је положио директно су тематски повезани са предложеном темом докторске дисертације.

Преглед положених испита

Назив предмета	Оцена	ЕСПБ	Датум
Пројектовање база података	(девет) 9	9	11.09.2015
Вештачка интелигенција и експертски системи	(десет) 10	9	04.02.2015
Принципи програмских језика	(девет) 9	9	02.09.2015
Интернет програмирање	(десет) 10	9	10.02.2015
Софтверски системи за рад у реалном времену	(десет) 10	9	11.02.2015
Одабрана поглавља из архитектуре рачунарских система	(десет) 10	9	22.01.2016
Сигурност и заштита рачунарских система	(девет) 9	9	20.01.2016
Енглески језик (десет)	(девет) 9	6	13.06.2016
Виртуелна стварност	(девет) 9	9	15.06.2016
Организација система дискова	(десет) 10	9	20.01.2016

Преглед одрађених обавеза

Назив обавезе	ЕСПБ
Научно стручни рад	3
Студијски истраживачки рад I	15
Студијски истраживачки рад II	15

Научноистраживачки рад

Кандидат је почео са израдом докторске тезе под менторством професорке др Сања Вранеш у фебруару 2016 на истраживањима у вези метода и окружења за побољшање квалитета описа великих количина линкованих података (*Linked Data datasets*). Циљ истраживања је био идентификација критеријума који су специфични за оцењивање квалитета отворених

података (ресурса) објављених на арапском језику и израда модел и архитектуру софтверског сервиса за анализу квалитета великих количина линкованих података на бази позадинског знања. Студијски истраживачки рад кандидата се превасходно базирао на истраживању технологија и иновативних метода представљања и обраде података из фармацеутске индустрије. Резултати кандидата су објављивањи у зборницима међународних научних скупова (категоризација М33).

Радови публиковани у зборницима међународних научних скупова:

- **Guma Abdulkhader Lakshen**, Sanja Vraneš, Valentina Janev: *Big data and quality: A literature review*. Proceedings of the 24th Telecommunications Forum (TELFOR 2016), 2016, pp. 802-805. DOI=[10.1109/telfor.2016.7818902](https://doi.org/10.1109/telfor.2016.7818902)
- **Guma Abdulkhader Lakshen**, Valentina Janev, Sanja Vraneš: *Challenges in Quality Assessment of Arabic DBpedia*. Proceedings of the 8th International Conference on Intelligence, Mining and Semantics (WIMS 2018), Article No. 15. DOI:[10.1145/3227609.3227675](https://doi.org/10.1145/3227609.3227675)
- **Guma Abdulkhader Lakshen**, Sanja Vraneš, Valentina Janev: *Linking Open Drug Data: the Arabic dataset*. Proceedings of the 9th International Conference on Information Society Technology (ICIST 2019), Kopaonik, Serbia, March 10 – March 13, 2019.
- **Guma Abdulkhader Lakshen**, Sanja Vraneš, Valentina Janev: *Quality Issues of Open Big Data Ecosystems: Toward Solution Development*. Proceedings of the 9th International Conference on Information Society Technology (ICIST 2019), Kopaonik, Serbia, March 10 – March 13, 2019.
- **Guma Abdulkhader Lakshen**, Valentina Janev, Sanja Vraneš: *Linking Open Drug Data: Lessons Learned*. Proceedings of the 18th International Conference on Computer Information Systems and Industrial Management Applications (CISIM 2019), Belgrade, September 19-21, 2019.

Учесће на пројектима

Кандидат је учествовао у реализацији неколико комерцијалних пројеката *Maintenance Management System MMS, a computerized maintenance management system*, LISCO (1989-1993); *Marketing Management Application. MMA, sales distribution system based on geographical and inhabitance of Libya*, GMPMSCO (2001-2003); *Designing a human resources management system HRMS*, GMPMSCO (2002-2005); *Classifying, monitoring, surveillance, of production equipment's and activities at Al-rabta Pharmaceutical factory*, as required by OPCW (2002-2011); *Study the requirements of unifying a computerized system at Al-Jabal Al-Gharbe University* (2011-2012).

Одбрана приступног рада за израду дисертације

Кандидат **Гума Абдулкадер Лакшен** је 11. јуна 2019. године полагао докторски испит на Електротехничком факултету у Београду, при чему је комисија у саставу:

1. др Бошко Николић, редовни професор, Електротехнички факултет у Београду
2. др Никола Томашевић, научни сарадник, Институт Михајло Пупин
3. др Милан Бјелица, ванредни професор, Електротехнички факултет

једногласно закључила да су кандидат и тема подобни за израду докторске дисертације на Електротехничком факултету Универзитета у Београду.

1.3. Оцена подобности кандидата за рад на предложеној теми

На основу претходног приказа досадашњег рада и оствареног искуства могу се сумирати следећи резултати:

- Током свог научноистраживачког рада, кандидат Гума Абдулкадер Лакшен се бавио проблемима који су у блиској вези са предложеном темом докторске дисертације.
- Кандидат Гума Абдулкадер Лакшен је до сада публиковао пет радова на конференцијама међународног значаја, док у рецензији су три рада за часопис.
- Кандидат Гума Абдулкадер Лакшен поседује одговарајуће знање и искуство потребно за решавање проблема из области предложене теме докторске дисертације.
- Кандидат Гума Абдулкадер Лакшен је положио све испите на докторским студијама Електротехничког факултета са просечном оценом 9,5. Поред тога, кандидат је пред Комисијом формираном од стране Наставно-научног већа Електротехничког факултета, на јавној усменој одбрани, успешно положио докторски испит, одржан 11.06.2019. године.

На основу наведених резултата, наша је процена да је кандидат Гума Абдулкадер Лакшен у стању да успешно реализује и заврши израду докторске дисертације.

2. Предмет и циљ истраживања

Предмет предложене докторске дисертације јесте развој методологије и софтверског решења за оцену квалитета података у апликацијама за рад са великим (*Big*) и интерлинкованим (*Linked*) подацима.

Ефикасно процесирање и управљање великим (*Big*) подацима се сматра за један од предуслова повећања конкурентске предности предузећа, узимајући у обзир да обим података свакодневно расте како у оквиру корпорацијских система, тако и на Вебу. У литератури се предвиђа да ће чак 33 зетабајта корисних података постојати на Вебу до краја ове године, а да ће тај број досећи 175 зетабајта до 2025. године. Организације из домена као што су здравство, медицинска / фармацеутска производња, енергетика и процесна индустрија, које граде своју основну делатност и услуге на способности за интеграцију података из различитих извора и имплементацију аналитичких сервиса, сматрају обраду великих и интерлинкованих података најпре и све више као императивну потребу за опстанак на тржишту.

Тема предложене докторске дисертације кандидата Гуме Лакшена усмерена је на задатак да детаљно истражи питања квалитета великих и повезаних (*Linked*) екосистема података, узимајући у обзир могућност поновног коришћења отворених података и, са овим у вези, теза дисертације ће предложити решење специфично за фармацеутску индустрију. Циљ је побољшати квалитет процеса интеграције отворених података у постојећем корпорацијском систему, осмислити методологију за развој апликације за рад са повезаним (*Linked*) подацима и омогућити претраживање тако генерисаног скупа података. Рад је мотивисан потребом да се омогући истраживачима из арапских земаља да употребом семантичких веб технологија повежу своје податке са отвореним подацима, као нпр. *DBpedia*-јом.

Проблем квалитета интерлинкованих података није досада дискутован у контексту методологије за рад са интерлинкованим подацима из Арапских земаља. У научној литератури уопште нема пуно радова који се баве приступима обраде интерлинкованих података, нпр. процесом генерисања, повезивања, објављивања и коришћења повезаних података. Једна од првих методологија повезаних података развијена је у Европском истраживачком пројекту LOD2 (2010-2014)[1] који је углавном био посвећен процесу отварања података и објављивања скупова података у машински читљивом формату. Методологије W3C конзорцијума [2] и других аутора као што су Хилард [3], Хаузенблас [4] и Вилазон-Теразас [5] се односе на податке из јавног сектора. У вези података из фармацеутске индустрије, Јовановић и Трајанов [6] предложили су нову методологију која

даје смернице компанијама о изради генеричких компоненти за поновну употребу у облику алата и сервиса за дати домен (нпр. управљање лековима). Завери [7] и Радуловић [8] су предложили *Linked Data Quality Model* и извршили тестовање са *DBpedia*-јом. Почетне инструкције за прорачун неких индикатора за квалитет интерлинкованих података је дат у раду од Завери [7], међутим њихова веза са имплементацијом неке *Linked Data* апликације за кориснике из арапског света није досада дискутована и проучавана у научној литератури.

Истраживачки рад у оквиру израде дисертације кандидата делом биће базиран на следећој литератури:

- [1] Auer, S., et al.: Managing the Life-Cycle of Linked Data with the LOD2 Stack. In: The Semantic Web-ISWC 2012. Boston: Springer Berlin Heidelberg, pp. 1–16, 2012.
- [2] W3C, Best Practices for Publishing Linked Data, <http://www.w3.org/TR/ld-bp/> (2016), last accessed 2019/05/01.
- [3] Hyland, B., Wood D.: The Joy of Data: A Cookbook for Publishing Linked Government Data on the Web. In: Linking Government Data, New York: Springer New York, 2011, pp. 3–26.
- [4] Hausenblas, M.: Linked Data Life Cycles, 2016. <http://www.slideshare.net/mediasemanticweb/linked-data-life-cycles>.
- [5] Villazón-Terrazas, et al.: Methodological Guidelines for Publishing Government Linked Data. In: Wood, D. (ed.) Linking Government Data, Ch. 2, 2011.
- [6] Jovanovic, M., Trajanov D.: Consolidating drug data on a global scale using linked data. *Journal of Biomedical Semantics*, 8(3), 2017.
- [7] Zaveri, A., Rula, A., Maurino, A., Pietrobon, R., Lehmann, J. Auer, S.: Quality assessment for linked data: A survey. *Semantic Web – Interoperability, Usability, Applicability*, Vol. 7, No. 1, 63-93. DOI: <http://dx.doi.org/10.3233/SW-150175> (2016)
- [8] Radulović, F., Mihindukulasooriya, N., García-Castro, R., Gómez-Pérez, A.: A Comprehensive Quality Model for Linked Data. *Semantic Web – Interoperability, Usability, Applicability*, Vol. 9, No. 1 (2018), (2018), 3-24, Special issue on Quality Management of Semantic Web Assets (Data, Services and Systems). DOI: <https://doi.org/10.3233/SW-170267> (2018)

3. Полазне хипотезе

У раду се полази од следећих хипотеза:

- Могуће је идентификовати главне разлоге за квалитет отворених података и предложити модел за опис индикатора квалитете интерлинкованих података.
- Квалитетни податаци са отворених извора омогућавају коришћење екстерног знања за обогаћивање релевантних база података предузећа и побољшање поузданости расуђивања у оквиру комплексних информационих система, као што су системи фармацеутске индустрије.
- Користећи стандардне речнике за опис података могуће је извршити интеграцију знања о лековима на глобалном нивоу и имплементирати систем за проверу квалитета интерлинкованих и публикованих података у *Linked Open Data* формату.
- Нови приступ публиковања података у *Linked Open Data* формату олакшава развој интероперабилних сервиса, доприноси стандардизацији размене података. и отвара нове могућности и хоризонте за расуђивање у дистрибуираним системима, као што су екосистеми великих података (*Big Data Ecosystem*).

4. Научне методе истраживања

У истраживањима у оквиру дисертације примењиваће се више различитих метода. Метода дескрипције ће бити коришћена за описивање функционалности система за проверу квалитета отворених података пре него се они повежу са подацима компаније. Метода класификације ће бити коришћена за разврставање анализираних софтверских алата за оцену квалитета података. Системска анализа ће се користити за дефинисање једног свеобухватног система за управљање интерлинкованих података у фармацеутској индустрији, као и његових компоненти/елемената и процеса.

Развој једног таквог система је интерактиван процес који се изводи у више фаза коришћењем различитих метода за сваку фазу, укључујући визуелно моделовање (*UML - Unified Modeling Language*), методу брзог развоја (*RAD -Rapid Application Development*), итд. Кандидат ће анализирати скупове података из различитих арапских земаља и развити приступ за трансформацију истих у *Linked Data* формату. Пројектовање и развој графичког интерфејса биће урађено технологијама отвореног кода.

Посебно, статистички метод биће коришћен за израчунавање индикатора квалитета за интегрисане и интерлинковане податке.

5. Очекивани научни допринос

- У оквиру докторске дисертације предложиће се нов приступ за интеграцију података у *Linked Data* формату узимајући у обзир специфичности арапског језика и потребу за провером квалитета података у свакој фази обраде и коришћења интерлинкованих података. Методологија ће бити примењена на отворене податке (класификовани са 2 звезде према моделу Тим Бернерс-Лее), а као резултат објавиће се скуп консолидованих података у *Linked Open Data* формату (класификовани са 5 звезда према моделу Тим Бернерс-Лее). Тако **главни допринос** истраживача је **консолидација података о лековима** из Сирије, Ирака, Либана и Саудијске Арабије, узимајући у обзир (релевантне) проблеме и аспекте у вези квалитета отворених података.
- Биће развијен оригинални софтверски сервис за процену квалитета интерлинкованих података узимајући у обзир разноликост (*variety*), варијабилност (*variability*), истинитост (*veracity*), ваљаност (*validity*) и димензију вредност (*value*). Семантички модел који се користи за опис и израчунавање индикатора квалитета интерлинкованих података ће **допринети стандардизацији сервиса у екосистема великих података** (*Big Data Ecosystems*).
- Концентришући се на питања квалитета података као што су тачност (*accuracy*), доследност (*consistency*) и релевантност (*relevancy*), у оквиру докторске дисертације, показаће се неке могућности примене отворених података (*DrugBank*, *DBpedia*) за обogaћивање база података о лековима једне компаније. У овом смислу, рад **доприноси спецификацији одговарајуће архитектуре семантичких интероперабилних пословних информационих система** у домену фармацеутске индустрије.
- Коначно, биће предложен и оригинални софтверски систем за ефикасно претраживање података о лековима (као што су цена, произвођачи, доступност земље, итд) из арапских земаља, као и проналажење релевантних информација на арапском језику за одређени лек из других идентификованих скупова података, као што су *DrugBank* и *Dbpedia*. Докторска дисертација ће тако дати **допринос у развоју нових генерација мултилингвалних софтверских апликација** за расуђивање, претраживање, откривање, и аутоматску испоруку знања у фармацеутској индустрији.

6. План истраживања и структура рада

- Истраживања на бази доступне иностране и домаће литературе из наведене области.
- Преглед најновијих истраживања из наведене области, као и тренутног стања наведене области.
- Рад на семантичком моделу за чување података.
- Рад на софтверском систему за оцену квалитета интерлинкованих података.
- Рад на софтверском модулу за израчунавање индикатора квалитета.
- Верификација очекиваних научних доприноса објављивањем резултата истраживања у научним часописима са SCI листе, као и у другим часописима и на конференцијама националног, регионалног и међународног значаја.
- Израда докторске дисертације на основу добијених резултата истраживања и научних доприноса.

7. Закључак и предлог

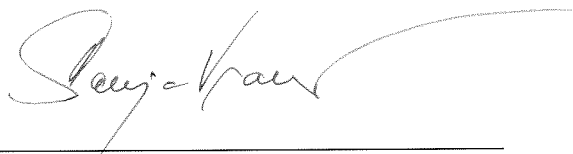
На основу приложене документације, као и непосредног увида у досадашњи научни рад кандидата Гуме Абдулкадер Лакшена, дипл.инж., Комисија је закључила да:

1. Кандидат испуњава све формалне и суштинске услове прописане Законом о високом образовању и Статутом Електротехничког факултета Универзитета у Београду за одобрење теме за израду докторске дисертације.
2. Предложена тема докторске дисертације кандидата Гуме Абдулкадер Лакшена „Окружење за анализу и оцену квалитета великих и повезаних података“ је у тренду са глобалним научним кретањима из области примене информационих и комуникационих технологија, чији се резултати тренутно објављују у истакнутим међународним часописима. Предложена тема је поред теоријског значаја, значајна и интересантна из угла праксе. Предложена истраживања у оквиру дисертације имају велики потенцијал да доведу до нових научних резултата.
3. Кандидат Гума Абдулкадер Лакшен се током свог научноистраживачког рада бавио тематиком која спада у најужу област дисертације. Комисија сматра да континуитет досадашњег научноистраживачког рада несумњиво указује на чињеницу да је кандидат подобан и способан за реализацију предвиђених циљева истраживања.
4. Кандидату се одобрава писање дисертације на енглеском језику уз наведени наслов „Окружење за анализу и оцену квалитета великих и повезаних података“ односно *“A framework for analysis and quality assessment of big and linked data”*.

Научна област дисертације је Електротехника и рачунарство, а ужа научна област је Рачунарска техника и информатика, за коју је Универзитет у Београду – Електротехнички факултет матичан. Имајући све претходно у виду, Комисија предлаже Наставно-научном већу Електротехничког факултета Универзитета у Београду да кандидату Гуми Абдулкадер Лакшену одобри израду докторске дисертације под горе наведеним називом. За ментора дисертације предлаже се др Сања Вранеш, редовни професор, Универзитет у Београду – Електротехнички факултет. Уз овај извештај прилаже се списак радова ментора на одговарајућем формулару Универзитета.

Београд, 25.06.2019.

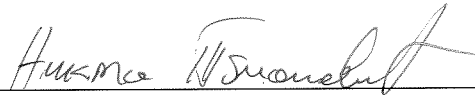
ЧЛАНОВИ КОМИСИЈЕ



др Сања Вранеш, редовни професор
Универзитет у Београду – Електротехнички факултет
(ментор)



др Бошко Николић, редовни професор
Универзитет у Београду – Електротехнички факултет



др Никола Томашевић, научни сарадник
Институт Михајло Пупин



др Милан Бјелица, ванредни професор
Универзитет у Београду – Електротехнички факултет