



**НАСТАВНО-НАУЧНОМ ВЕЋУ
ПРИРОДНО-МАТЕМАТИЧКОГ ФАКУЛТЕТА
И ВЕЋУ ЗА ПРИРОДНО-МАТЕМАТИЧКЕ НАУКЕ
УНИВЕРЗИТЕТА У КРАГУЈЕВЦУ**

На седници Наставно-научног већа Природно-математичког факултета у Крагујевцу, одржаној 25.04.2018. године (број одлуке: 290/XIII-I), предложени смо, а на седници Већа за природно-математичке науке одржаној 16.05.2018. године (број одлуке: IV-01-352/19), изабрани смо за чланове Комисије за подношење извештаја за оцену научне заснованости теме и испуњености услова кандидата за израду докторске дисертације под насловом: „Биоинформатичка платформа за извршавање Federated SPARQL упита над онтолошким базама податка и детектовање сличних података утврђивањем њихове семантичке повезаности“ кандидата Марије Ђокић – Петровић, студента докторских академских студија рачунарских наука. Комисија је у следећем саставу:

1. др Мирјана Ивановић, редовни професор, председник комисије
Природно-математички факултет, Универзитет у Новом Саду
Ужа научна област: Рачунарске науке и информатика
Датум избора у звање: 29. IV 2002
2. др Ненад Стефановић, ванредни професор, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Информациони системи
Датум избора у звање: 10. XI 2016
3. др Снежана Марковић, доцент, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Физиологија животиња и човека и молекуларна биологија
Датум избора у звање: 14. IX 2016
4. др Татјана Стојановић, доцент, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Вештачка интелигенција

Датум избора у звање: 15. VI 2016

5. др Владимир Цвјетковић, доцент, предложени ментор рада
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Информатика у физици
Датум избора у звање: 12. I 2016

На основу увида у достављени материјал Комисија подноси следећи

ИЗВЕШТАЈ

1. Научни приступ проблему предложеног нацрта докторске дисертације и процена научног доприноса крајњег исхода рада

Биоинформатика је један од резултата изузетног напретка информатичке технологије чији је значај упоредив са индустријском револуцијом и масовним коришћењем разних облика енергије. Развој база података у другој половини претходног века, пораст моћи процесирања, и интеграција рачунарских мрежа омогућили су примену информатичких технологија и у биолошким истраживањима. Непосредни значај биоинформатике данас, као интердисциплинарне области, базира се на огромном броју биолошких података који се адекватно могу користити и процесирати једино применом савремених информатичких технологија. Резултати биоинформатике доприносе побољшаној дијагностици болести, откривању генетске предиспозиције болести, омогућавају дизајнирање нових лекова, итд. Оно што је од кључног значаја у домену биоинформатике данас, јесте доступност потребних података битних за истраживања, као и информација о томе да такви подаци негде већ постоје. Да би истраживачи у домену биоинформатике могли лако да изводе закључке и доносе конкретне одлуке неопходне за даља истраживања, они морају бити у стању да на брз, ефикасан, и једноставан начин приступе свим потребним подацима. Битан предуслов за то је да су потребни подаци јавно доступни, повезани и да су развијени механизми за претраживање и приступ. Многе базе података су се развијале у изолованом окружењу, без поштовања основних стандарда биоинформатичке заједнице. Ове хетерогене базе података, које су део многих независних и високо специјализованих

ресурса, често користе различите формате, конвенције и речнике за представљање својих података. Са друге стране, многи комплексни и мулти-дисциплинарни биоинформатички пројекти спроводе истраживања над високо дистрибуираним подацима и њихова реализација и успешност у великој мери зависе од доступности и ефикасне употребе ових података. Додатно, многи извори података се преклапају и покривају (прикривају) сличне податке, стварајући на тај начин хомогене или полу-хомогене изворе података. У таквим случајевима семантички однос оваквих скупова података често је нејасан и неопходно је применити одговарајуће методе за анализу података (енгл. *Data Mining*), како би се открили слични подаци.

У циљу решавања наведених проблема биоинформатичка заједница користи технологије семантичког веба. У том погледу развијени су многи семантички репозиторијуми и софтверска решења, који су знатно допринели истраживачким активностима на биоинформатичкој сцени. Такође, разне научно-истраживачке организације, институције и лабораторије прате актуелне биоинформатичке трендове и дају допринос развоју биоинформатичке заједнице као и развоју нових технологија и софтверских решења. CPCTAS (Centre for PreClinical Testing of Active Substances је назив FP7 пројекта GA206809) Лабораторија за ћелијску и молекуларну биологију (у даљем тексту Лабораторија) као део Института за биологију и екологију Природно-математичког факултета Универзитета у Крагујевцу бави се преклиничким тестирањима биоактивних супстанци и развија сопствени информациони систем за подршку истраживањима у сарадњи са Институтом за математику и информатику са истог Факултета, који обухвата и биоинформатички аспект истраживања. Једна од основних области истраживања Лабораторије је преклиничко тестирање активних супстанци, потенцијалних лекова за канцер. Поље рада Лабораторије припада једном од значајних подпроцеса у оквиру важне биоинформатичке гране - релационом дизајнирању и откривању лекова. Биоинформатичка платформа (у даљем тексту платформа), која је предмет дисертације, развијена је за потребе Лабораторије и има за циљ да биоинформатичка истраживања учини продуктивнијим и ефикаснијим.

2. **Образложење предмета, метода и циља који уверљиво упућује да је предложена тема од значаја за развој науке**

Предмет дисертације је развој и усавршавање функционалности биоинформатичке софтверске платформе, која је резултат вишегодишње сарадње кандидата са Лабораторијом на развоју информационог система за подршку истраживањима у оквиру кога је развијана биоинформатичка веб платформа. Основне функционалности веб платформе јесу извршавање Federated SPARQL упита над иницијалним и кориснички одабраним биоинформатичким изворима података, као и детектовање сличних података, које је засновано на утврђивању њихове семантичке повезаности.

Извршавање Federated SPARQL упита изводи се над базама података које користе Resource Description Framework (RDF) као модел података. Упити су дизајнирани у складу са потребама Лабораторије, али су од генералне важности и за друга биоинформатичка истраживања. Резултати извршених упита се могу додатно филтрирати и тиме допринети побољшању њихове релевантности. Филтрирање подразумева одабир специфичних својстава (предиката) приликом динамичке пројекције RDF структуре скупа података и извршавање динамички генерисаних star-shaped SPARQL упита.

Циљ дисертације је подршка и унапређивање истраживања у разним областима биологије коришћењем развијене биоинформатичке платформе базиране на технологијама семантичког веба за откривање релевантних и комплементарних информација потребних у биолошким истраживањима. Аутоматизованим претраживањем више различитих извора биоинформатичких података организованих као семантички репозиторијуми, директно се користи огроман потенцијал ових извора података на начин који истраживачи дефинишу сагласно специфичним потребама својих истраживања. Конкретно, могуће је открити биолошке тест системе – биолошке мете (нпр., специфичне имортализоване ћелијске линије канцера) који су се користили за експериментална истраживања и дефинисање молекуларних механизма биолошког деловања одговарајућег активне супстанце (потенцијални лек). Ови биолошки тест системи могу припадати различитим базама података. Даље се уз помоћ опције филтрирања могу издвојити „повољне“ биолошке мете, односно оне које су коришћене у успешно спроведеним експериментима. Успешност експеримента се процењује на основу цитотоксичности одређене активне супстанце, а одређује помоћу IC50 вредности (половина максималне инхибиторне концентрације тестиране активне супстанце). У ширем смислу, овај квантитативни показатељ је индикатор мере/дозе у којој је тестирана супстанца/лек (инхибитор) неопходна да се

инхибира дати биолошки процес за половину. Биолошке мете које имају атрибут „повољних“ могу се користити за будућа експериментална истраживања. Шанса успешности извођења таквих експеримената сигурно је већа од оних са насумично одабраним биолошким тест системима.

Друга значајна функционалност платформе јесте детектовање сличних података, која је заснована на утврђивању њихове семантичке повезаности. Алгоритам (у даљем тексту алгоритам) који је развијен за потребе ове функционалности представља нов и јединствен приступ за детекцију сличних података над онтолошким базама података. Користе се принципи *Text-Data Mining*-а, модел векторског простора за математичку репрезентацију података и косинусну меру сличности за нумеричко одређивање сличности података. Одређивање сличних података свакако може утицати на планирање будућих експеримената у домену биоинформатике. Конкретно, када се утврди група „повољних“ биолошких мета, применом алгоритма могу се издвојити оне биолошке мете које су сличне, иако оне могу припадати различитим базама података, односно имати различите *Uniform Resource Identifier (URI)* спецификације. На тај начин се може сузити претрага „повољних“ биолошких мета и уштедети ресурси који су неопходни за извођење будућих експеримената. Развијени алгоритам се може применити не само у домену биоинформатике, него и у било којој другој области која покрива онтолошко моделовање података.

Методологија развоја софтвера примењена у дисертацији полази од основних постојећих решења прихваћених од стране широке заједнице истраживача како оних који раде на развоју, тако и корисника, и сагледавања потреба и могућности за даљим унапређењима. Постојеће решење од кога се полази је базирано на коришћењу технологија семантичког веба које се у истраживањима везаним за дисертацију даље усавршава развојем веб платформе за комбиновни приступ и претраживање независно развијаних различитих извора података који су од интереса, базираних на основним принципима технологија семантичког веба. Такав приступ такође отвара проблем проналажења различито означених истих или сличних података за шта је у дисертацији развијено опште решење у виду алгоритма за аутоматизовано откривање таквих података.

3. Образложење теме за изradу докторске дисертације које омогућава закључак да је у питању оригинална идеја или оригиналан начин анализирања проблема

Полазну основу дисертације чине досадашња истраживања примене технологија семантичког веба у домену биоинформатике. Основне хипотезе истраживања су да технологије семантичког веба:

- нуде богате и добро дефинисане моделе за представљање и интеграцију података
- омогућавају агрегацију хетерогених података коришћењем експлицитне семантике
- омогућавају поновну употребу података
- олакшавају претрагу података

Развој метаподатака у области биологије је важан за примену технологија семантичког веба у домену биоинформатике, јер омогућава семантичку интеграцију података коришћењем RDF оквира као основе за семантичко представљање података. Додатно, онтологије као формални модели за експлицитно представљање концепата и односа између њих, користе се за решавање питања хетерогености у изворима података. Упитни језик SPARQL (SPARQL Protocol and RDF Query Language), као стандардни упитни језик за RDF, омогућава извршавање глобалних упита над различитим изворима података који могу бити дистрибуирани и третирани као једна велика семантичка база података на вебу. Генерално, технологије семантичког веба олакшавају процес имплементације, интеграције, поновне употребе и претраге података у домену биоинформатике. Њихова примена олакшава откривање релевантних и комплементарних података, над којима се могу применити разне методе за анализу података (*Data Mining*), што доприноси квалитетнијим истраживањима у домену биоинформатике.

У дисертацији ће бити представљене и детаљно образложене фазе истраживања које су претходиле развоју биоинформатичке платформе, која је предмет дисертације. Централна компонента биоинформатичке платформе је онтологија названа PIBAS (Преклиничка Испитивања БиоАктивних Супстанци је назив Пројекта Министарства за просвету, науку и технолошки развој Републике Србије, бр. ИИИ41010) чија је намена моделирање структуре варијетета експеримената у CPSTAS лабораторији или некој сличној лабораторији дефинисањем веза и односа између кључних концепата који карактеришу разноврсне експерименталне процедуре. У овом кораку извршена је анализа неких

постојећих онтологија које садрже концепт експеримента, поређење са PIBAS онтологијом, усвајање и прилагођавање одговарајућих концепата што је било могуће с обзиром на карактеристику транспарентности PIBAS онтологије (у даљем тексту онтологија). Онтологија је развијена коришћењем OWL синтаксе, а за имплементацију је коришћен онтолошки едитор Protégé. Дизајном је предвиђено да се подаци појединачних експеримената складиште у посебне онтолошке фајлове а да се њиховом композицијом креира онтолошка база података названа CPCTAS (по називу Лабораторије). Развијен је и прототип софтвера, који омогућава претрагу експеримената CPCTAS базе података (у даљем тексту база података) по различитим критеријумима који се могу комбиновати на различите начине према потребама истраживача.

Следећа фаза истраживања је интеграција базе података са другим биоинформатичким решењима. Извршена је компарација онтологије са онтологијама других биоинформатичких база података, које припадају глобалним семантичким репозиторијумима као што су Bio2RDF, Chem2Bio2RDF и EMBL-EBI. Компарациони поступак је извршен коришћењем OWLDiff алата (енгл. *Protégé plug-in*). Интеграција је резултовала креирањем нових класа и својстава онтологије, које су омогућиле повезивање инстанци базе података са инстанцама екстерних база. Извршена интеграција је обезбедила извођење различитих Federated SPARQL упита, који су углавном креирани у складу са истраживачким активностима Лабораторије и који се извршавају на платформи. Предефинисани упити се извршавају једноставном селекцијом иницијалних критеријума претраге, а изводе се над базама података глобалних репозиторијума Bio2RDF, Chem2Bio2RDF, EMBL-EBI, и развијене базе података, али и над кориснички одабраним скуповима података. У циљу представљања ове функционалности спроведено је детаљно истраживање SPARQL синтаксе и образложени су основни типови SPARQL упита. Базе података наведених репозиторијума који се користе за извођење иницијалних SPARQL упита на платформи биће детаљно представљене и анализирани. Биће приказано како се добијени резултати на платформи могу додатно побољшати на основу пројекције RDF структуре скупа података, селекцијом одговарајућих концепата (предиката) и извршавањем star-shaped SPARQL упита. Резултати добијени извршавањем упита биће анализирани у сарадњи са истраживачима сродних Лабораторија, како би се утврдила њихова релевантност у планирању будућих експеримената.

У дисертацији ће детаљно бити анализирана кључна функционалност платформе - могућност детекције сличних података. Ова функционалност се базира на утврђивању семантичке повезаности података, користећи принципе *Text-Data Mining*-а, примену модела векторског простора као и косинусне мере сличности. Алгоритам, који је имплементиран у циљу детекције сличних података, примењује се на резултате, односно URI (*Uniform Resource Identifier*) спецификације, добијене извршавањем Federated SPARQL упита. Алгоритам ће за сваку URI спецификацију у почетном скупу података, креирати одговарајуће SELECT SPARQL упите и на основу селекције предиката прикупити објекте који имају форму стрингова (текста). У овом кораку тежи се ка аутоматској фаворизацији предиката. Над текстуалним вредностима ће се затим применити основне процедуре кластеризације текста (енгл. *Text clustering*). Резултати кластеризације ће бити представљени коришћењем модела векторског простора. У циљу утврђивања коначних резултата, као нумеричка мера сличности између вектора, користиће се косинусна сличност која је непосредно преузета из теорије проналажења информација. У складу са имплементираном функционалношћу биће спроведено детаљно истраживање из области технологија семантичког веба; вршиће се анализа и поређење термина семантичке сличности (енгл. *Semantic similarity*) и семантичке повезаности података (енгл. *Semantic relatedness*); детаљно ће бити образложени принципи *Text-Data Mining*-а, укључујући основне процедуре кластеризације текста. Један део истраживања биће базиран на примени различитих мера сличности над бинарним подацима (векторима), како би се утврдила најбоља мера сличности (у овом случају косинусна мера сличности). Резултати добијени применом алгоритма биће анализирани у сарадњи са истраживачима Лабораторије, ради утврђивања валидности алгоритма.

4. Усклађеност дефиниције предмета истраживања, основних појмова, предложене хипотезе, извора података, метода анализе са критеријумима науке уз поштовање научних принципа у изради коначне верзије докторске дисертације

У уводном делу дисертације изложиће се основни концепт биоинформатике као и значај у домену савременог истраживања. У овој целини укратко би биле представљене неке од најзначајнијих грана биоинформатике са акцентом на базе податка и процес дизајнирања и

откривања лекова. Овај део дисертације укључивао би анализу литературних података и утврђивање чињеничних стања и проблема са којима се биоинформатичка заједница суочава. Такође, био би представљен концепт веба, хронолошки развој веба као и утицај на домен биоинформатике. Пажња би највише била усмерена на потенцијал који технологије семантичког веба нуде у циљу решавања биоинформатичких проблема (имплементације, интеграције и претраге података).

Тема другог поглавља дисертације је трећа генерација веба - семантички веб. У складу са тим извршило би се представљање архитектуре семантичког веба, а затим преглед и анализа технологија, које се данас активно промовишу у домену биоинформатике. Овај део дисертације би могао послужити и као водич за примену технологија семантичког веба у различитим доменима.

Треће поглавље дисертације представиће биоинформатичку област истраживања CPCTAS лабораторије - преклиничко тестирање активних супстанци. Детаљно би била представљена структура „Експеримент“ концепта и били би објашњени неки од кључних термина ове структуре. Биће дат приказ истраживања која су спроведена за потребе Лабораторије, а која су реализована применом технологија семантичког веба: конструкција као и развој онтологије (онтологија Експеримента), развој онтолошке базе података; функционалност прототипа софтвера за претрагу и интеграцију базе података.

Централни део дисертације биће посвећен опису функционалности биоинформатичке платформе и обухватаће неколико поглавља.

У четвртом поглављу ће бити детаљно представљена архитектура платформе и основне компоненте.

Основне функционалности платформе биле би фокус наредног, петог поглавља:

- Извршавање предефинисаних Federated SPARQL упита над иницијалним базама података: Овај део дисертације би укључивао примере извођења различитих Federated SPARQL упита, који су дизајнирани на основу истраживачких активности Лабораторије. Кроз тестне сценарије би се представио значај и намена предефинисаних упита. У складу са тим детаљно би биле представљене базе података, који су делови великих репозиторијума (Chem2Bio2RDF, Bio2RDF и EMBL-EBI), а које се користе за дефинисање упита. Акцент би био усмерен на начин на који се примењују технологије семантичког веба у циљу превазилажења актуелних биоинформатичких проблема, као и компарација са базом

података, која је такође коришћена као иницијална база података. Биће укључен компаративни преглед литературе актуелних софтверских решења за извођење Federated SPARQL упита у домену биоинформатике, и преглед литературе која се бави идентификацијом релевантних извора података (енгл. *Source selection*) неопходних за генерисање Federated SPARQL упита.

- Извршавање Federated SPARQL упита над иницијалним и кориснички одабраним онтолошким базама података: У овом случају биоинформатичка платформа би нудила опцију трајног складиштења података кориснички одабраних извора, чиме би се омогућила њихова поновна употреба. Овај део дисертације бавио би се и утврђивањем таксономије концепата кориснички одабраних база неопходних за проширење иницијалних упита. Циљ ове функционалности је да омогући лабораторијама и научно-истраживачким институцијама да промовишу интерне податке и да укажу на свој значај у домену биоинформатике.
- Динамичко филтрирање резултата упита: Ова функционалност има за циљ побољшање релевантности резултата добијених извршавањем предефинисаних упита. Она омогућава одабир скупова података које се користе у упиту, пројекцију њихових RDF структура и селекцију одговарајућих својстава (предиката) на основу којих ће се генерисати star-shaped SPARQL упити. Извршавањем ових упита могу се добити додатне информације, које могу утицати на ток планираних биоинформатичких истраживања.

Напредне функционалности платформе - детектовање сличних података, представиће се у шестом поглављу. Један део овог поглавља био би посвећен анализи литературних података са циљем утврђивања чињеничних стања овог проблема у домену биоинформатике. Била би анализирана и рефлексија овог проблема у резултатима који су настали као последица извршавања предефинисаних упита. Представиће се одговарајућа решења и методе које омогућавају детекцију сличних података, и њихова компарација са процесом детекције сличних података из дисертације. Детаљно ће се образложити алгоритам развијен за потребе функционалности платформе, и преглед одговарајућих *Data-Text Mining* принципа и математичких метода коришћених за његову имплементацију. Алгоритам ће се применити над резултатима добијених извршавањем Federated SPARQL упита и биће изведена серија тестних сценарија. У циљу верификације алгоритма, резултати ће бити анализирани у сарадњи са истраживачима у Лабораторији.

Очекује се да резултати добијени у овој докторској дисертацији значајно допринесу процесу откривању знања које утиче на планирање будућих биоинформатичких истраживања. Тиме се стварају услови да наредна истраживања и експерименти буду ефикаснији и продуктивнији.

5. Предложени ментор докторске дисертације и комисија

За ментора ове дисертације предлаже се доцент др Владимир Цвјетковић са Природно-математичког факултета Универзитета у Крагујевцу, ужа научна област информатика у физици, под чијим руководством је кандидат отпочео са упознавањем семантичког веба као и са применама и истраживањима у области биоинформатике. Доцент Цвјетковић и кандидат су више година радили на развоју информационог система за подршку истраживањима Лабораторије, чија је компонента развијена биоинформатичка платформа која је предмет дисертације. Предложени ментор је аутор или коаутор 11 радова са SCI листе који обухватају области експертних система, примена веб технологија на мерење и управљање физичким системима као и представљање и управљање знањем.

Компетенције чланова комисије

Предложени ментор и чланови комисије имају значајно академско научно и наставно искуство и својим научним компетенцијама покривају интердисциплинарну област ове докторске дисертације. Професор др Мирјана Ивановић је експерт у области рачунарских наука и информатике. Професор др Ненад Стефановић је експерт у области информационих система. Доцент др Татјана Стојановић је експерт у домену вештачке интелигенције. Доцент др Снежана Марковић је експерт у области биолошких наука, ужа област молекуларна биологија и физиологија које су везане за истраживања подржана развијеном биоинформатичком платформом.

6. Подобност кандидата

Марија В. Ђокић-Петровић је рођена 11. фебруара 1986. године у Крагујевцу. Основну школу „Карађорђе“ завршила је у Рачи 2001. године, као носилац дипломе „Вук Караџић“. Другу крагујевачку гимназију, општи смер, завршила је 2005. године као ученик генерације и носилац дипломе „Вук Караџић“. На Природно-математички факултет у

Крагујевцу, група Математика, смер математика-информатика, уписала се школске 2005/06, где је и дипломирала маја 2010. године са просечном оценом 8,92. Током студија учествовала је на EDIT летњој школи фирме ComTrade ITSS, Београд, Србија, а затим је у истој фирми обавила стручну праксу у трајању од два месеца, радећи на тестирању софтвера. Током студија била је корисник студентског кредита Министарства просвете и спорта, а последње године студија и стипендиста општине Рача. Школске 2010/2011 године уписује докторске студије из области рачунарских наука на Природно-математичком факултету у Крагујевцу, где је и положила све испите предвиђене програмом докторских студија са просечном оценом 9,29.

Од маја 2011. године до јула 2017. године радила је као истраживач-приправник и истраживач-сарадник за ужу научну област Програмирање на Институту за Математику и информатику, Природно-математичког факултета, Универзитета у Крагујевцу. У том периоду је била ангажована на пројекту Министарства „Преклиничка Испитивања БиоАктивних Супстанци“ (евиденциони број пројекта П41010, руководилац пројекта др. Снежана Марковић, доцент Природно-математичког факултета, Универзитета у Крагујевцу), где је била члан информатичког тима, који се бави администрацијом, одржавањем и развојем веб базираног информационог система за потребе истраживачког Центра. Ту је започела рад и истраживање у области семантичког веба са применом у биоинформатици и сарадњу са истраживачима из Лабораторије. Била је ангажована и на реализацији вежби из предмета ”Основи програмирања” на студијском програму за стицање стручног/академског звања Дипломирани математичар и Дипломирани физичар. Своју професионалну каријеру наставила је у аустријским фирмама Virtual World Services GmbH из Беча и mything GmbH из Граца, преко којих је ангажована на пројекту Техничког Универзитета у Грацу, Аустрија. Рад на овом престижном Универзитету омогућио јој је усавршавање знања из области програмирања и научно-истраживачког рада.

Члан је уредништва сајта Austria-forum.org, који својим квалитетним садржајем, у виду многих биографија и публикација, представља значајан допринос аустријском друштву.

Члан је управног одбора аустријског удружења ”NEU Graz“ (Verein für Neoösterreicher und EU-Bürger Graz), где својом информатичком подршком доприноси развоју веб платформе удружења.

Активни је члан Друштва информатичара Србије. Њено ангажовање на Техничком Универзитету у Грацу у Аустрији покренуло је сарадњу овог Универзитета са Друштвом информатичара Србије.

Као члан управног одбора удружења „NEU Graz“ активно се бави хуманитарним радом. Често помаже удружења и активисте који се боре за заштиту животиња.

Говори енглески и немачки језик.

7. Преглед научно-истраживачког рада кандидата.

Радови на SCI листи М 20 категорије

1 **Djokic-Petrovic, M.**, Cvjetkovic, V., Yang, J., Zivanovic, M., & Wild, D. J. (2017). PIBAS FedSPARQL: a web-based platform for integration and exploration of bioinformatics datasets. *Journal of biomedical semantics*, 8(1), 42, ISSN 2041-1480, DOI 10.1186/s13326-017-0151-z, **M22 IF 1.845, 2016**

2 **Djokic-Petrovic, M.**, Pritchard, D., Ivanovic, M. and Cvjetkovic, V. (2016), IMI Python: Upgraded CS Circles web-based Python course. *Computer Applications & Engineering Education*, 24(3), pp. 464 - 480, ISSN: 1099-0542, 24: 464–480, doi:10.1002/cae.21724, **M23 IF 0.694, 2016**

Радови М 50 категорије

3 Cvjetkovic, V., **Djokic, M.**, Arsic, B., and Curcic, M. (2014). The ontology supported intelligent system for experiment search in the scientific research center. *Kragujevac Journal of Science*, (36), 95-110. ISSN: 1450-9636, **M53**

Радови М 30 категорије

4 V. M. Cvjetkovic, **M. Djokic**, B. Arsic, Owl based modelling and visualisation of arbitrary semantic data structure, Proceedings of the 5th International Conference "Science and Higher Education in Function of Sustainable Development", Business Technical College, pp. 2:13-19, Uzice, October 04 - 05, 2012, <http://sed.vpts.edu.rs/> (ISBN 978-86-83573-26-4, **M33**)

5 V. Cvjetkovic, **M. Djokic**, B. Arsic, Ontology Visualization, Proceedings of the 1st Virtual International Conference on Advanced Research in Scientific Areas (ARSA-2012), vol. 1, issue 1, pp. 1999--2004, Slovakia, December 3 - 7, 2012, <http://www.arsa-conf.com> (ISBN: 978-80-554-0606-0, ISSN: 1338-9831, **M33**)

6 V. Cvjetkovic, **M. Djokic**, B. Arsic, Wikipedia Browsing with DBpedia, Proceedings in EIIC - The 2nd Electronic International Interdisciplinary Conference, vol. 2, issue 1, pp. 470--475, 2013, <http://www.eiic.cz/> (ISBN: 978-80-554-0762-3, ISSN: 1338-7871, **M33**)

7 V. Cvjetković, **M. Djokic**, B. Arsić, Semantically based customized search on local web site, Proceedings of the 1st Virtual International Conference on Advanced Research in Scientific Areas (ARSA-2012) vol. 2, issue 1, pp. 453--458, Slovakia, December 3 - 7, 2013, <http://www.arsa-conf.com> (ISBN: 978-80-554-0825-5, ISSN: 1338-9831, **M33**)

8 **M. Djokic**, N. Stefanovic, Application of Semantic Web in tourism information systems, Proceedings of the 3rd International Conference on Information Society Technology (ICIST-2013), pp. 3:130-135 Kopaonik, March 3 - 6, 2013, <http://e-drustvo.org/icist/2013/default.html> (ISBN: 978-86-85525-12-4, **M33**)

9 B. Arsic, **M. Djokic**, N. Stefanovic, Mapping ebXML standards to ontology, Proceedings of the 4th International Conference on Information Society Technology (ICIST-2014), vol. 1, pp. 198-203, Kopaonik, Serbia, March 9 - 13, 2014 (ISBN: 978-86-85525-14-8, **M33**)

10 B. Arsic, **M. Djokic**, V. Cvjetkovic, P. Spalevic, M. Zivanovic and M. Mladenovic, Integration of bioactive substances data for preclinical testing with Cheminformatics and Bioinformatics resources, Proceedings of the 23rd International Electrotechnical and Computer Science Conference, vol. 1, issue 1, pp. 146-149, ERK, September 22 - 24, 2014, Portorož, Slovenia (ISSN:1581-4572, **M33**)

11 B. Arsic, **M. Djokic**, V. Cvjetkovic, P. Spalevic and S. Ilic, Semantic search framework for distributed semantically based cheminformatics and bioinformatics datasets, 5th International Conference on Information Society and Technology (ICIST 2015), Society for Information Systems and Computer Networks, pp. 518 - 522, Serbia, 8. - 11. Mar, 2015 (ISBN: 978-86-85525-16-2, **M33**)

12 V. Cvjetkovic, **M. Djokic**, Semantic web based organization of scientific bibliography references, The 3rd International Virtual Conference on Advanced Scientific Results (SCIECONF-2015), EDIS - Publishing Institution of the University of Zilina, vol. 1, no. 3, pp. 230 - 235, Slovakia, 25. - 29. May, 2015, pp. 230 - 235, 2015. (ISBN: 978-80-554-0891-0, ISSN: 1339-9071, **M33**)

Радови М 80 категорије

13 V. Cvjetkovic, S. Markovic, B. Arsic, J. Zizic, **M. Djokic**, Семантички базирана кастомизирана претрага на локалном веб сајту: <http://cpctas-lcmb.pmf.kg.ac.rs>, Универзитет у Крагујевцу, Природно-математички факултет, 2013. **M84**

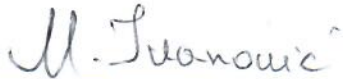
ЗАКЉУЧАК

Кандидат Марија Ђокић – Петровић је студенткиња докторских академских студија рачунарских наука на Институту за математику и информатику Природно-математичког факултета Универзитета у Крагујевцу где је положила све програмом предвиђене испите са просечном оценом 9,29. У току докторских студија објавила је 2 рада из М20 категорије (један рад М22 и један рад М23), један рад категорије М53, има 9 саопштења на међународним конференцијама штампаним у целини категорије М33 и једно техничко решење категорије М84, што је укупно 13 библиографских референци. Наведена библиографија указује да се кандидат бави актуелном научно заснованом проблематиком и истраживањима из области семантичког веба и да је сходно томе предложена тема актуелна и научно заснована, као и да кандидат има способности да успешно уради предложену дисертацију.

На основу свега изложеног Комисија закључује да је предложена докторска дисертација са називом „Биоинформатичка платформа за извршавање Federated SPARQL упита над онтолошким базама податка и детектовање сличних података утврђивањем њихове семантичке повезаности“ кандидата Марије Ђокић – Петровић, дипломираног

математичара, студенткиње докторских академских студија рачунарских наука, научно значајна и оригинална и да кандидат испуњава услове за израду дисертације, те предлаже Наставно-научном већу Природно-математичког факултета и Већу за природно-математичке науке Универзитета у Крагујевцу да јој се одобри рад на предложеној дисертацији, под руководством предложеног ментора, др Владимира Цвјетковића са Природно-математичког факултета Универзитета у Крагујевцу.

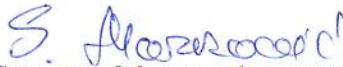
Комисија



1. др Мирјана Ивановић, редовни професор, председник комисије
Природно-математички факултет, Универзитет у Новом Саду
Ужа научна област: Рачунарске науке и информатика



2. др Ненад Стефановић, ванредни професор, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Информациони системи



3. др Снежана Марковић, доцент, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Физиологија животиња и човека и молекуларна биологија



4. др Татјана Стојановић, доцент, члан комисије
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Вештачка интелигенција



5. др Владимир Цвјетковић, доцент, предложени ментор дисертације
Природно-математички факултет, Универзитет у Крагујевцу
Ужа научна област: Информатика у физици