

УНИВЕРЗИТЕТ У БЕОГРАДУ

ФАКУЛТЕТ ОРГАНИЗАЦИОНИХ НАУКА

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Тијане Чомић

Одлуком Наставно-научног већа Факултета организационих наука 3/59-8 од 07.06.2017 именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Тијане Чомић, под насловом:

УНАПРЕЂЕЊЕ ЗВАНИЧНЕ СТАТИСТИКЕ ПРИМЕНОМ *BIG DATA* КОНЦЕПТА

После прегледа достављене Дисертације и других пратећих материјала и разговора са Кандидатом, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Кандидат Тијана Чомић уписала је мастер студије, опција: економетрија, школске 2008/09. године, на Економском факултет Универзитета у Београду. Положила је све планом и програмом предвиђене испите и 2009. године одбранила мастер рад под насловом „Одређивање маргиналних ефеката фактора сиромаштва применом модела са квалитативном зависном променљивом“. Од 2013. уписана је на Факултету организационих наука Универзитета у Београду, као студент докторских студија на смеру Операциона истраживања.

Докторске студије је уписала 2013/2014 године. Након положених испита према наставном плану и програму, пријавила је приступни рад под називом „Унапређење званичне статистике променом *Big data* концепта“. Комисија за преглед и одбрану приступног рада и оцену научне заснованости пријављене докторске дисертације је формирана 30.12.2015. године, Одлука број 3/4-6. Приступни рад је одбранила пред Комисијом дана 20.12.2016. године.

Одлука о усвајању извештаја Комисије о научној заснованости пријављене докторске дисертације донета је 15.03.2017. године, одлука бр. 3/27-6. Одлуком Универзитета у Београду бр. 61206-1231/2-17 од 27.03.2017. године дата је сагласност на предлог теме докторске дисертације Тијана Чомић под називом „Унапређење званичне статистике применом *Big Data* концепта“. Одлуком Наставно-научног већа Факултета организационих наука 3/35-4 од 12.04.2017. године одобрена је израда докторске дисертације, а за ментора је именован др Зоран Радојичић, ред. проф. Факултета организационих наука. Ментор је 24.10.2018. године известио Наставно-научно веће Факултета организационих наука да је Тијана Чомић завршила израду докторске дисертације. Наставно-научно веће Факултета организационих наука је именovalo Комисију за оцену завршене докторске дисертације 31.10.2018. године, одлука бр. 3/189-4.

1.2. Научна област дисертације

Корисници званичних статистичких података, истраживачи, креатори политика и остали, све више користе могућности за употребу мноштва података који не долазе из извора званичне статистике и претварају их у њима корисне информације. Осим доступности, ове изворе података карактерише и брзина дисеминације. Без обзира на јасно одсуство методолошких поставки, које могу довести (и доводе) у питање квалитет добијених података, корисници наводе да им тако добијени подаци могу бити од користи за брзо идентификовање проблема, потреба, пружање услуга, али и за предвиђање и спрачавање криза, а ради добробити становништва. Основни проблем који се разматра у дисертацији је истраживање потенцијала примене *Big Data* концепта у домену званичне статистике, која се базира на веома чврстим законодавним и методолошким оквирима. Предмет истраживања докторске дисертације је развој *Big Data* модела процеса статистичког истраживања који представља унапређене традиционалног модела, пре свега у домену практичне имплементације и унапређења методолошког поступка. Предложени модел је конципиран тако да буде апликативан у систему званичне статистике у Републици Србији са могућношћу проширења у другим истраживачким доменима.

Према предмету истраживања и коришћеној методологији дисертација припада техничким наукама, прецизније се може сврстати у ужу научну област рачунарска статистика за коју је Факултет организационих наука Универзитета у Београду матичан.

Ментор др Зоран Радојичић, ред. проф., поседује одговарајуће компетенције за вођење дисертације у виду низа објављених научних радова у светским часописима, као и учешћу у већем броју домаћих и међународних пројеката из области статистике, званичне статистике и статистичких истраживања.

1.3. Биографски подаци о кандидату

Тијана Чомић (рођ. Милојевић) рођена је у Крагујевцу 8. септембра 1984. Основну школу и Прву крагујевачку гимназију завршила је у Крагујевцу. Уписала је 2003. године Економски факултет Универзитета у Београду, где је дипломирала 2007. године на смеру Статистика, информатика и квантитативна анализа. На истом факултету завршила је и мастер студије, опција: економетрија. Мастер рад на тему „Одређивање маргиналних ефеката фактора сиромаштва применом модела са квалитативном зависном променљивом“ одбранила је 2009. године са оценом 10. Од 2013. уписана је на Факултету организационих наука Универзитета у Београду, као студент докторских студија на смеру Операциона истраживања.

Од 2007. до 2016. године била је запослена у Републичком заводу за статистику на позицији статистичара-аналитичара. Била је руководиоца групе за Анкету о приходима и условима живота.

У сарадњи са Бечким инстиутом радила је као део тима који је истраживао проблем неједнакости и јавних политика. Рад је презентован на неколико међународних скупова организованих од стране GDN. Од стране УНИЦЕФ-а била је ангажована на додатној обради података MICS-а. Тренутно је ангажована на пројекту NetSILC3 који има за циљ унапређење методологије EU-SILC-а, на нивоу Европске уније. Током 2017. године била је ангажована од стране регионалне канцеларије УНИЦЕФ-а, као консултат на пословима упоређивања европских истраживања са MICS истраживањем у циљу употпуњавања доступности индикатор одрживог развоја. Тренутно је од стране регионалне канцеларије УНИЦЕФ-а ангажована као регионални консултат MICS-а за истраживања на домаћинствима.

Од децембра 2016. ради као независни експерт у области статистичког праћења индикатора одрживог развоја и животног стандарда, на пројектима које спроводи УНИЦЕФ, Европска комисија, UNDP, итд.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Дисертација је обима од 203 странице, садржи 40 слика и графичких приказа, 14 табела, 260 литературних навода, прилог, биографију и изјаве аутора.

Структура докторске дисертације обухвата следеће целине: уводна разматрања, технике и методе прикупљања и обраде података у званичној статистици, *Big Data* концепт, примена концепта *Big Data* у званичној статистици, развој методологије статистичког истраживања на бази *Big Data* концепта, основне поставке *Big Data* модела, примена модела – студија случаја: индикатори квалитета живота. Следе научни стручни доприноси, будући правци истраживања и закључна разматрања.

Докторска дисертација се састоји из следећих поглавља и потпоглавља:

1 Увод

- 1.1 Предмет истраживања
- 1.2 Циљ истраживања
- 1.3 Полазне хипотезе рада
- 1.4 Научне методе истраживања

2 Технике и методе прикупљања и обраде података у званичној статистици

- 2.1 Развој званичне статистике
- 2.2 Извори података
- 2.3 Методи прикупљања података

3 Big Data концепт

- 3.1 Big Data и информатичка револуција
- 3.2 Big Data, основне дефиниције
- 3.3 Друштвени медији

4 Примена концепта Big Data у званичној статистици

- 4.1 Нове технике прикупљања података
- 4.2 Big Data у УНЕЦЕ
- 4.3 Big Data у ESS

5 Развој методологије статистичког истраживања на бази Big Data концепта

- 5.1 Основне напомене
- 5.2 Основне претпоставке за реализацију модела
 - 5.2.1 Законодавни и стратешки оквир за примену Big Data у званичној статистици
 - 5.2.2 Приступ подацима и партнерства
 - 5.2.3 Кадрови
 - 5.2.4 Технолошка инфраструктура

6 Основне поставке Big Data модела

- 6.1 Фазе животног циклуса Big Data пројекта
- 6.2 Фаза 1: Утврђивање потреба
- 6.3 Фаза 2: Пројектовање
- 6.4 Фаза 3: Реализација и тестирање производног система
- 6.5 Фаза 4: Прикупљање података
- 6.6 Фаза 5: Обрада података
- 6.7 Фаза 6: Анализа података
- 6.8 Фаза 7. Евалуација
- 6.9 Фаза 8. Дисеминација
- 6.10 Фаза 9. Архивирање

7 Примена модела – студија случаја: индикатори одрживог развоја

- 7.1 Фаза 1: Утврђивање потреба
 - 7.1.1 Утврђивање потреба за информацијама
 - 7.1.2 Консултације и потврђивање потреба и дефинисање потребних резултата
 - 7.1.3 Идентификовање концепата (статистичких стандарда)
 - 7.1.1 Провера расположивости података
- 7.2 Фаза 2. Пројектовање

7.2.1	Процена ризика извођења Big Data пројекта
7.2.2	Пројектовање резултата и дефинисање променљивих
7.2.3	Избор метода и инструмената прикупљања података
7.2.4	Методологија статистичке обраде
7.3	Анализа потенцијала за примену циљане методе истраживања друштвених медија у систему званичне статистике Републике Србије
7.4	Сентимент анализа
7.5	Закључна разматрања студије случаја
8	Научни и стручни доприноси
8.1	Научни доприноси
8.2	Стручни доприноси
8.3	Друштвени допринос
9	Будућа истраживања
10	Закључак
11	Референтна литература
12	Списак слика
13	Списак табела

2.2. Кратак приказ појединачних поглавља

Прво поглавље представља увод у истраживање у којем је дефинисан предмет истраживања, циљеви дисертације, полазне хипотезе и методе истраживања.

У другом поглављу дате су основне дефиниције статистике као науке и званичне статистике из које је она настала. Приказан је правни основ за функционисање, као и начин функционисања, почев од националног нивоа, преко нивоа Европске уније, до Статистичке комисије Уједињених нација. Дат је кратак осврт на методологију традиционалних статистичких истраживања са акцентом на методе које се користе у Републичком заводу за статистику Србије.

У трећем поглављу су дефинисани основни појмови и концепти *Big Data*. Наводи се неколико одабраних дефиниција, које су у основи сличне и базирају се на компилацији података и информација, толико великих и комплексних да је тешко спровести њихову обраду и анализу помоћу стандардних, тренутно доступних метода и алата. На основу анализе Европског статистичког система, приказана је листа доступних *Big Data* извора по доменима и институцијама земаља који су их идентификовали. Посебна пажња усмерена је на нивое структурираности податка и могућностима употребе недовољно структурираних информација, будућих да већина података која се користи у истраживањима на бази *Big Data* концепта је дато у неструктурираној форми.

У четвртном поглављу приказани су основни резултати истраживања који се односе на примену концепта *Big Data* у званичној статистици. Имајући у виду комплексност концепта *Big Data* и специфична знања и регулативе неопходне за његову имплементацију у систем званичне статистике, на нивоу Европске уније одлучено је да се уложи заједнички напори свих националних статистика како би се заједничким радом тестирале могућности коришћења овог концепта. У овом поглављу се наводе и анализирају примери коришћења или планови за коришћење и тестирање *Big Data* у

званичној статистици. На основу спроведене анализе закључак је да алтернативни извори података, због свог волумена и нивоа структурираности, и даље представљају велику непознаницу за традиционалну статистику. Из тог разлога, мора се прво извршити темељна анализа и процена квалитета података доступних из *Big Data* извора

Предмет петог поглавља представља развој методологије статистичког истраживања на бази *Big Data* концепта. Приказан је еволутивни развој методолошког поступка у сфери статистичких истраживања, пратећи еволутивни развој науке, према датим парадигмама. Указано је на најважније концептуалне разлике у приступу истраживањима заснованим на подацима (*Big Data* приступ) и истраживањима базираним на тестирању хипотеза (традиционални приступ). Као основна предност, уједно и мана *Big Data* приступа, наводи се да процес изградње модела заснован на великој количини података у мањој мери зависи од теоријских претпоставки и ограничења. Имајући у виду предмет истраживања дисертације, а узевши у обзир основне принципе система званичне статистике, имплементација концепта *Big Data* у процес статистичких истраживања императивно се базира на научним принципима и методама. Пратећи дефинисане циљеве истраживања, идентификовани су најважнији изазови и дефинисане основне претпоставке за развој методологије која омогућава имплементацију предложеног решења - *Big Data* модела процеса статистичког истраживања (*Big Data* модел). У складу са наведеним изазовима, дефинисани су предуслови за развој методологије која омогућава имплементацију *Big Data* модела са посебним освртом на Републику Србију, који се огледају у дефинисању законодавног и стратешког оквира, приступа подацима, потенцијалним партнерствима, увидом у кадровску функцију и захтеваној технолошкој инфраструктури.

У шестом поглављу изнете су основне поставке *Big Data* модела. У циљу хармонизације статистичких процеса у домену званичне статистике, захтев је био да *Big Data* модел, што је могуће више, прати фазе традиционалног модела процеса статистичког истраживања. Дат је преглед дефинисаних фаза, почев од утврђивања потреба, пројектовања, преко реализације и тестирање производног система, које укључују: прикупљање, обраду и анализа података, закључно са евалуацијом, дисеминацијом и фазом архивирања. Акценат је стављен на подфазе *Big Data* модела које се значајно разликују од традиционалног модела истраживања. Између осталог, истакнута је неопходност процене ризика у сфери коришћења података о личности и нарушавања приватности. У ту сврху је дефинисан је инструмент у форми контролне листе за процену ризика извођења *Big Data* пројекта (Фаза 2: Пројектовање, подфаза 2.1).

У складу са постављеним циљевима истраживања докторске дисертације, у седмом поглављу су приказани основни резултати тестирања предложеног методолошког поступка. Елаборирана је студија случаја која се бави редеофинисањем (и имплементацијом) индикатора квалитета живота на националном нивоу у Републици Србији. Овај индикатор омогућава дефинисање (и имплементацију) УН индикатора 1.4.1. „Удео становништва које живи у домаћинствима која имају приступ основним услугама“. У студији су описане релеванте фазе предложеног *Big Data* модела за спроведено истраживање. На основу анализе предлога појединачних индикатора и потенцијалних извора података, закључено је да се поједини индикатори могу се

пратити методом истраживања друштвених медија и одговарајућом *Big Data* аналитиком. Такође, извршена је анализа потенцијала за примену циљане методе и одговарајућих техника истраживања друштвених медија (сентимент анализа) у систему званичне статистике Републике Србије. Добијени резултати су потврдили полазне хипотезе и указали на могућности и практична ограничења у имплементацији *Big Data* модела у производњи статистичких резултата у домену званичне статистике.

У осмом поглављу дат је преглед научних, стручних и друштвених доприноса дисертације. Будући правци истраживања приказани су у деветом поглављу. У Закључку је дат преглед постигнутих резултата и научних доприноса дисертације. Списак литературе садржи релевантне референце за област дисертације.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1 Савременост и оригиналност

Предмет дисертације је значајан и припада актуелним областима истраживања могућности примене савремених информационо-комуникационих технологија и пословне аналитике у области, не само званичне статистике, већ и статистичких истраживања, уопште. Велики број научних и стручних часописа, конференција, књига, ресурса са веба и заинтересованост светских истраживачких и консултатских кућа, као и самих истраживача, говори у прилог актуелности шире теме докторске дисертације.

Истраживања у домену примене *Big Data* у званичне статистике су иницирана и подржана са највишег нивоа, од стране Статистичке комисије Уједињених нација и Европске комисије (Евростата). Увидом у обимну литературу може се потврдити да је *Big Data* концепт нашао упориште у водећим статистичким институцијама и организацијама широм света. Покренути су многи пројекти у циљу подршке имплементације још већег броја, не само пилот пројеката, у овој области. На нивоу ЕУ направљено је јединствена платформа: CEF Big Data Test Infrastructure (BDTI), као комплетно решење које егзистира у виртуелном окружењу и омогућава европским организацијама (укључујући и статистичке заводе) да експериментишу с технологијама *Big Data* технологијама. Осим саме инфраструктуре, BDTI регистрованим корисницима ставља на располагање одређене скупове података и аналитичка решења.

Докторска дисертација детаљно анализира досадашња истраживања и сазнања из области примене *Big Data* у области статистичких истраживања и системима званичне статистике. Област производње званичне статистике карактерише чврсто методолошко утемељење базирано на међународним стандардима и препорукама у којима квалитет података има доминантну улогу. Резултати се публикују у строго дефинисаној динамици, која захтева доста времена, од момента дефинисања истраживања до саме дисеминације. Практично, истраживања се спроводе прикупљањем података из примарних извора (анкете и пописи) или из секундарних извора (регистри). Без обзира на значајне резултате који су постигнути у области примене рачунарски подржаних анкетних истраживања (CATI, CAPI и CAWI) до сада није забележен озбиљнији помак у развоју моделског, општег приступа истраживањима базираним на *Big Data* концептима. У дисертације се указује да овај

концепт има потенцијал да значајно унапреди процес прикупљања, складиштења, обраде, анализе и представљања података различитих структура из мноштва извора. Унапређења се првенствено односе на приступ алтернативним изворима података, који омогућавају производњу нових статистичких индикатора, брже извештавање корисника и растерећење даваоца података. Нарочито је истакнута могућност подршке креаторима политика у случајевима доношења одлука које захтева брзу реакцију, као што су економске кризе.

Да би остварила ова концепција, у дисертацији се сагледава неопходност дефинисања моделског приступа истраживањима базираним на *Big Data* концептима. Као једна од кључних препрека за пуну имплементацију овог концепта евидентиран је проблем недостатка савремене методологије, нарочито у сфери узорка, статистичког оцењивања и закључивања.

Након детаљне анализе, у дисертацији је предложено једно системско решење кроз пројектовање и имплементацију *Big Data* модела за истраживања у домену званичне статистике, које се може применити и шире, у свим областима статистичких истраживања (медицинским, маркетиншким, истраживањима јавног мњења и сл.).

Ужи предмет истраживања ове дисертације се односи на унапређење методолошких и организационих поступака статистичких истраживања са аспекта примене савремених информационих технологија, нарочито *Big Data*. Осим тога, дисертација се бави анализом претпостављених предуслова неопходних за потпуну имплементацију предложеног решења.

Најважнији резултат истраживања ове докторске дисертације је развој оригиналног модела у чијој основи је *Big Data* модел процеса статистичког истраживања који представља унапређени статистички пословни модел (GSBPM) и представља његову модификацију у циљу имплементације *Big Data* концепта. Предложени модел не садржи појединачна технолошка решења и алате, већ је базиран на ближем методолошком одређењу. Приказане су и анализиране фазе животног циклуса истраживања *Big Data* модела. Приказано решење се базира на имплементацији Контролне листа за процену ризика извођења *Big Data* пројекта у домену коришћења података о личности и

У складу са Правилником о поступку провере оригиналности докторских дисертација које се бране на Универзитету у Београду, Универзитетска библиотека Светозар Марковић је извршила проверу оригиналности дисертације коришћењем програма iThenticate и у потпуности је потврђена оригиналност дисертације. Оригиналност и неопходност у приступу решавања проблема и добијених резултата у оквиру ове дисертације потврђују и радови који су публиковани у међународном часопису са импакт фактором категорије **M23** и саопштени на међународним научним скуповима (**M33**).

На основу изложеног, може се закључити да добијени резултати докторске дисертације представљају научни допринос у односу на постојеће стање, и отварају простор за даља истраживања. Значај теме докторске дисертације огледа се и у чињеници да су резултати истраживања примењиви у пракси на велики број различитих области, које карактерише проблематика статистичких истраживања.

3.2 Осврт на референтну и коришћену литературу

Последњих година вршена су интензивна истраживања у области *Big Data* и њихове примене у различитим сферама истраживања. Многобројни резултати публиковани су у часописима, књигама, презентовани на разним конференцијама и од стране водећих светских компанија и организација. Кандидат је приликом писања докторске дисертације користио обимну и актуелну литературу, која обухвата 307 референци. Литература обухвата значајне ауторе и радове из релевантних научних часописа са SCie и SSCI листе, радове са конференција. У складу са темом и областима којима се дисертација бави, у раду су коришћене и публикације водећих консултантских кућа (Ernst & Young, Forreseter, Gartner, McKinsey, и компанија (Google, IBM, Oracle, SAS) и званичне публикације међународних организација и институција (Уједињене нације, OECD, Светска банка, Међународни монетарни фонд, статистички заводи и институти, итд.).

Неки од коришћених извора су:

1. Abbasi, A., A. Hassan, M. Dhar (2014). Benchmarking Twitter Sentiment Analysis Tools, Conference: 9th Language Resources and Evaluation Conference, Reykjavik, Iceland,
2. Bok, B., D. Caratelli, D. Giannone., A. M. Sbordon, A. Tambalotti (2018). Macroeconomic Nowcasting and Forecasting with Big Data, Annual Review of Economics, Volume 10, <https://doi.org/10.1146/annurev-economics-080217-053214>
3. Cambria, E., B. Schuller, B. Liu, H. Wang, C. Havas (2013). Statistical Approaches to Concept-Level Sentiment Analysis, IEEE Intelligent Systems, Volume: 28, Issue: 3, May-June 2013
4. Cavallo, A, R. Rigobon (2016). The Billion Prices Project: Using Online Prices for Measurement and Research. Journal of Economic Perspectives 30.2 (2016): 151–178.
5. Cavanillas, J. M., E. Curry, W. Wahlster, Editors (2013), New Horizons for a Data-Driven Economy A Roadmap for Usage and Exploitation of Big Data in Europe, Springer Open, 2015
6. Daas, P.J.H., Puts, M.J. (2014). Social media sentiment and consumer confidence, Statistics paper series no 5 / september 2014, European Central Bank, <https://www.ecb.europa.eu/pub/pdf/scpsps/ecbsp5.en.pdf>
7. Daas, P.J.H., Puts, M.J., Buelens, B. and van den Hurk, P.A.M. (2015). Big Data as a Source for Official Statistics. Journal of Official Statistics 31(2), pp. 249-262. <http://dx.doi.org/10.1515/JOS-2015-0016>

8. Gandomi, A., M. Haider (2015). Beyond the hype: Big data concepts, methods, and analytics, *International Journal of Information Management*, Volume 35, Issue 2, April 2015, Pages 137-144, <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
9. Hackl, P. (2016). Big Data: What can official statistics expect?, *Statistical Journal of the IAOS* 32 (2016) 43–52 43, DOI 10.3233/SJI-160965, IOS Press
10. Hammer, C. L, D. C. Kostroch, G. Quirós, and STA Internal Group (2017). Big Data: Potential, Challenges, and Statistical Implication, IMF - Statistics Department, JEL Classification Numbers: E0, Y2, Z0, <https://www.imf.org/~media/Files/Publications/SDN/2017/sdn1706-bigdata.ashx>
11. Lyko, K. Nitzschke, M., Ngonga Ngomo, A-C (2016) Big Data Acquisition, chapter 4 in *New Horizons for a Data-Driven Economy, A Roadmap for Usage and Exploitation of Big Data in Europe*, Springer Open, доступно на: <https://link.springer.com/content/pdf/10.1007%2F978-3-319-21569-3.pdf>
12. Ospina, A.V. (2018). Big Data for resilience storybook: Experiences integrating Big Data into resilience programming. Winnipeg: International Institute for Sustainable Development. Retrieved from www.iisd.org
13. Samuel, S. J, R.V.P, Koundinya, K. Sashidhar, C.R. Bharathi (2015). A survey on big data and its research challenges, *ARPN Journal of Engineering and Applied Sciences*. Vol. 10, NO. 8, MAY 2015
14. Tam, S-M, Clarke, F. (2015). Big Data, Official Statistics and Some Initiatives by the Australian Bureau of Statistics, *International Statistical Review* (2015), 0, 0, 1–13 doi:10.1111/insr.12105
15. Vaccari, C. (2014). Big Data in Official Statistics, PhD Thesis, Università degli studi di camerino, School of advanced studies

За потребе анализе потенцијала за примену истраживања друштвених заједница и интернет истраживања уопште коришћене су публикације Републичког завода за статистику Србије и Регулаторне агенције за електронске комуникације и поштанске услуге – Ратела, док су код дефинисања индикатора одрживог развоја коришћени извори Тима за социјално укључивање и смањење сиромаштва, Владе Републике Србије и DevInfo Србија.

3.3. Опис и адекватност примењених научних метода

Научне методе, примењене у дисертацији, у потпуности одговарају предмету истраживања и захтевима у погледу провере постављених хипотеза. Основни метод истраживања базирао се на прикупљању, анализи и класификацији постојеће литературе и теоријских и експерименталних резултата у предметној области. За потребе истраживања, сакупљана је и проучавана научна и стручна литература и анализирани успешни примери из праксе која се односи на употребу *Big Data* концепта у статистичким истраживањима, са акцентом на званичну статистику. Доступни извори су детаљно анализирани и систематизовани, са циљем да се покаже оправданост и корисност развоја новог модела за имплементацију *Big Data* концепта у системима званичне статистике.

У циљу тестирања постављених хипотеза у току израде дисертације примењен је већи број научних метода. Примењене су следеће опште методе: методе дескриптивне анализе, методе индукције-дедукције, методе компаративне анализе, аналитичке и статистичке методе (методе мултиваријационе анализе, корелациона анализа, реузорковање). Поред општих метода истраживања, коришћене су и посебне методе које проистичу из специфично постављеног предмета и циља истраживања: традиционалне методе статистичких истраживања базиране на примарним и секундарним изворима података, као и методе за обраду и анализу *Big Data*. Могућност примене предложеног методолошког оквира је приказана на реалним примерима из праксе званичних статистичких институција. Верификација предложеног *Big Data* модела спроведена је анализом студије случаја која се односи дефинисање индикатора одрживог развоја у Републици Србији.

Резултати су презентовани текстуално, описивањем и приказани су кроз више табела, слика и дијаграма са упоредним резултатима. Истраживање је интердисциплинарно, јер укључује методологију, статистику, информатику и друге научне дисциплине.

На основу анализе докторске дисертације може се закључити, да примењене научне методе и технике, по свом значају и структури, одговарају теми дисертације и спроведеном истраживању.

3.4. Применљивост остварених резултата

С обзиром на актуелност теме и важност званичне статистике, применљивост добијених резултата истраживања је значајна. Резултати докторске дисертације указују на могућност имплементације предложеног модела у систем званичне статистике у Србији, како би се наставио процес хармонизације и одржао корак са најбољим статистичким институцијама у Европи и свету.

Такође, добијени резултати указују и на додатне могућности примене модела, уз одређене модификације, за друге области истраживања, као што су маркетиншка и социолошка истраживања, истраживања јавног мњења, медицинска истраживања и слично.

За будуће примене предложеног модела у Србији неопходно је испунити одређене предуслове, који су наведени у дисертацији, а у најкраћем тичу се стварања стратешког и правног оквира на националном нивоу, додатног отварања података, иницирању приватно-јавних-академских партнерствима у области *Big Data*, школовања одговарајућег кадра (статистичари, научници и стручњаци за податке) и рад на јачању технолошке инфраструктуре.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Области научног интересовања Тијане Чомић су: рачунарска статистика, статистичка анализа и званична статистика, пре свега у домену квантитативне анализе животног стандарда и услова живота, као и рањивих група. Стекла је значајно практично

искуство учествујући на већем броју пројеката у области истраживања у домену званичне статистике и анкетних истраживања у домаћинствима.

У току израде докторске дисертације, Тијана Чомић показала је способност за проналажење, упознавање и разумевање проблема истраживања са више аспеката на основу доступне литературе, као и креативног приступа његовом решавању. Учила је главне недостатке и проблеме постојећих решења из области статистичких истраживања, нарочито у домену званичне статистике и конципирала и спровела истраживање са циљем да се они превазиђу кроз израду системског решења базираног на моделском приступу. Такође, показала је потребну аналитичност и практичну оспособљеност за имплементацију предложеног модела што је верификовала публикавањем постигнутих резултата у научном часопису и конференцијама. Од укупно 9 радова које је у својој професионалној каријери до сада објавила, резултати истраживања проистекли из рада на докторској дисертацији објављени су у часопису категорије M22 и зборнику радова категорије M33. Сви објављени радови кандидата су из области примењених статистичких истраживања, што је у складу са предметом докторске дисертације.

На основу наведеног, потврђене су способности кандидата да адекватно интерпретира, уопштава и дискутује резултате истраживања и сматрамо да кандидат Тијана Чомић поседује потребно знање и искуство за самосталан научни рад.

4. ОСТВАРЕН НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Најзначајни научни доприноси дисертације су:

- Унапређење методолошког поступка за спровођење истраживања у домену званичне статистике, коришћењем *Big Data* концепта.
- Дефинисање *Big Data* модела процеса статистичког истраживања и
- Унапређење методологије за мониторинг циљева одрживог развоја.

Предмет истраживања докторске дисертације представља развој *Big Data* модела процеса статистичког истраживања који представља унапређене традиционалног модела, пре свега у домену методолошког поступка. Предложени модел је конципиран тако да је буде примењив у систему званичне статистике Републике Србије, са могућношћу проширења у другим истраживачким доменима.

Потенцијална имплементација предложеног методолошког поступка приказана је на моделу статистике одрживог развоја у Републици Србији, и то по фазама имплементације, закључно са анализом и евалуацијом основних предуслова. Посебна пажња усмерена је на процену ризика извођења *Big Data* пројекта у домену коришћења података о личности и нарушавања приватности.

Услед комплексности и осетљивости система званичне статистике, који се базира на чврстим методолошким упориштима, у складу са међународним статндардима и

препорукама, до сада није било много радова који су са методолошког аспекта истраживали ову област. У том смислу материја изложена у докторској дисертацији представља посебну вредност.

Један од научних доприноса дисертације представља унапређење методологије за мониторинг циљева одрживог развоја. Кроз студију случаја која обухвата Републику Србију извршена је анализа индикатора одрживог развоја, по УН и ЕУ приступу. За изабрани индикатор 1.4.1: „Удео становништва које живи у домаћинствима која имају приступ основним услугама“, предложен је методолошки поступак за дефинисање индикатора животног стандарда, применом дефинисаног *Big Data* модела. Имплементација предложеног решења омогућава да се унапреди процес мониторинга одрживог развоја у наредном периоду, пре свега у скраћеној периодици извештавања.

Стручни доприноси

Основни стручни доприноси дисертације се огледају у могућности примене предложене методологије у различитим областима које покрива званична статистика:

- национални рачуни
- економске статистике
- друштвене статистике.

Осим тога, предложена методологија истраживања се може користити и у другим областима, као што су: медицинска истраживања, истраживања јавног мњења, маркетиншка истраживања, истраживања друштвених медија и сл.

Анализа дефинисаних предуслова за развој методологије која омогућава имплементацију *Big Data* модела процеса статистичког истраживања у Републици Србији указује да је већина ових предуслова делимично испуњено.

За једну област *Big Data* – онлине истраживања, у Републици Србији, извршена је анализа потенцијала примене истраживања друштвених заједница и интернет истраживања уопште. Указано је на потенцијал, ограничења и очекивани правац развоја методологије у области, правећи аналогију са традиционалним телефонским истраживањима.

Са становишта друштвене корисности, резултати истраживања могу имати вишеструке импликације:

- Резултати истраживања помоћи ће да се анализира проблематика даљег унапређења техника и метода статистичких истраживања у домену званичне статистике.
- Резултати истраживања омогућавају да се анализира потенцијал примене *Big Data* концепта на нивоу целе управе и државне администрације у Републици Србији.
- Резултати истраживања указују на неопходне предуслове, пре свега законодавне и стратешке, који морају бити испуњени за потпуну имплементацију предложеног концепта.

- Резултати истраживања доприносе да се прецизније одреде потребни временски, технолшки, финансијски и људски ресурси, неопходни за увођење истраживања базираног на *Big Data* концепту у програм и план званичне статистике.
- Резултате истраживања, пре свега предложену методологију, могу користити и друге институције, привредна друштва и појединци који се баве статистичким истраживањима, како би успешно имплементирале истраживања базирана на *Big Data* концепту.

4.2. Критичка анализа резултата истраживања

Кандидат је у докторској дисертацији разматрао методолошки приступ и модел процеса статистичког истраживања. Истраживачки рад је заснован на постојећим научним резултатима и имплементираним решењима, са циљем представљања егзактних смерница које ће моћи да служе будућим корисницима које усвајају *Big Data* концепт у домену званичне статистике. Анализирана је обимна литература, као и различите методе и технике у области статистичких истраживања.

Резултат истраживања је оригинални модел процеса статистичког истраживања дефинисан за потребе званичне статистике. Модел предложен у овој дисертацији је флексибилан, проширив, пружа добре перформансе, омогућава интеграцију различитих извора података система који укључују и нестандартне формате података. За потребе тестирања примењивости модела на просторима Републике Србије, спроведено је истраживање које је потврдило полазне претпоставке и омогућило развој модела.

4.3. Верификација научних доприноса

Списак објављених радова

Категорија M22:

1. Cappa, C.; D. Mont, M. Loeb, C. Misunas, J. Madans, **T. Comic**,; de F. Castro (2018).The development and testing of a module on child functioning for identifying children with disabilities on surveys. III: Field testing, Disability and health journal, ISSN: 1876-7583, Vol: 11, Issue: 4, Page: 510-518, Publisher(s): Elsevier BV, Publication Year: 2018, PMID: 30049638, DOI: 10.1016/j.dhjo.2018.06.004 (IF 1.863; 5-Year Impact Factor: 2.129)

Категорија M33:

1. Jovičić, M., & **Milojević, T.** (2010). Poverty and inequality changes in Serbia as the result of global instability. The 11-th Bi-Annual Conference of the EACES, August 26-28, 2010, Comparing Responses to Global Instability. Tartu, Estonia. ISBN 978-9985-4-0613-7.
<http://ec.ut.ee/eaces2010/artiklid/Jovicic%20Milojevic-Poverty%20and%20inequality.pdf>

2. **Čomić, T.** (2016): Challenges and experiences in reporting on SDGs on people with disabilities - Serbian Perspective, International Seminar on Data for Sustainable Development Goals: Data Disaggregation, United Nations Statistics Division (UNSD), Statistics Korea (KOSTAT), Seoul, Republic of Korea, 3-4 November 2016, ESA/STAT/AC.324/7
3. **Čomić, T.**, A. Đoković, D. Vukmirović (2017): Potencijal primene Big Data koncepta u domenu zvanične statistike, Infoteh-Jahorina, Vol. 16, March 2017.
4. **Čomić, T.** (2018): Inclusion of production for own consumption in the household disposable income concept: Impact on the income distribution and on key EU income-based indicators, Net-SILC3 International Conference, Comparative EU Statistics on Income and Living Conditions, Athens, 19-20 April 2018

Категорија М40: Монографија националног значаја:

1. Jovičić M., **T. Milojević** (2009). Ocena efekata zamrzavanja penzija i zarada u javnom sektoru na siromaštvo, Ekonomska politika srbije u 2010. godini - Ka novom modelu makroekonomske stabilnosti, Redaktori: S. Stamenković i B. Živković. Naučno društvo ekonomista sa Akademijom ekonomskih nauka i Ekonomski fakultet u Beogradu, CIDEF. ISBN 978-86-403-1048-2
2. Stamenković, S., M. Kovačević, B. Živković, V. Vučković, **T. Milojević** (2011). Ekonomska politika srbije u 2011: Početak ili kraj razvojnog modela? Nova strategija razvoja privrede Srbije – Izazovi ekonomske politike u 2011. godini, Redaktori: Jurij Bajec и Miomir Jakšić, Naučno društvo ekonomista Srbije sa Akademijom ekonomskih nauka i Ekonomski fakultet u Beogradu

Посебне публикације:

1. Ivančev, O., M Jovičić and **T. Milojević** (2010). Income Inequality and Social Policy in Serbia, The Vienna Institute for International Economic Studies, Balkan Observatory Working Paper No. 86, October 2010
2. **Милојевић, Т.**, Ђ. Вуксановић (2009). Анализа сиромаштва - интеграција избеглих лица, Анализа карактеристика сиромаштва у Србији, Уредник: Јелена Марковић, Влада Републике Србије, Тим потпредседника Владе за имплементацију Стратегије за смањење сиромаштва

Пројекти:

1. IPA 2015 Multi-beneficiary statistical cooperation programme, Non-key expert for the Pilot project 7.2 Sustainable Development Goal (SDGs) indicators in selected countries in Western Balkan, GOPA. (2018 - on-going),
2. (on-going) UNICEF Regional Office, Geneva, regional MICS consultant, in charge for survey methodology. (2017 - on-going),

3. Household economic Survey, Kingdom of Saudi Arabia, consultant for data analysis and report writing, GOPA. (2017 - on-going),
4. UNICEF Kazakhstan, Review and critically assess work done for identification and collection of data sources used for the production of the Statistical Yearbook on children of Kazakhstan; Refine a list of core indicators for monitoring the situation of children in the country; Establish SDG baselines on selected core children-related SDG indicators at the national and disaggregated levels. For disaggregated levels, further analysis of the 2015-16 MICS as well as calculation of confidence intervals may be necessary; Provide support in user-friendly visualization and presentation of data and deliverables, helping to produce web-page on children. (2017).
5. UNICEF Turkmenistan, Consolidation, further analysis of MICS5, disaggregation of the existing data on children and women to serve as SDG baselines and inform the upcoming Situation Analysis of Children and Women in Turkmenistan, preparation of dashboards for regional analysis in Turkmenistan. (2017).
6. Uzbekistan National Nutrition Survey 2017, Technical support in the capacity building of national team on household mapping and listing, Uzbekistan, UNICEF CO. (2017).
7. Conducting baseline assessment for child-related SDG-s in the region; Enhance the potential for convergence between MICS and EU-SILC/LFS; Revision of MICS6 survey instruments for additional indicators; Support RO in revision of CSPs for different countries. (2016-2017).
8. Mapping of the baseline data availability for CEE/CIS Region for children related SDGs and enhance the potential for convergence between MICS and EU-SILC/LFS. (2016)
9. Revision of social inclusion indicators, Social Inclusion and Poverty Reduction Unit, Government of the Republic of Serbia (SIPRU) and UNICEF. (2016-2018).
10. Net-SILC3, Eurostat and LISER. (2016-2018).
11. Strengthening the Serbian Statistical System by upgrading methodologies and standards and by the appliance of good practice, Contract no: 371-594, IPA 2012 National. (2016 - on-going).
12. Field Testing of the UNICEF/WG Module on Child Functioning and Disability, SORS and UNICEF. (2015-2016).
13. Secondary analyses of data from the survey on income and living conditions (SILC), Social Inclusion and Poverty Reduction Unit, Government of the Republic of Serbia (SIPRU). (2015).
14. Additional Analysis of MISC data – Poverty, UNICEF. (2015).
15. MICS5 (Multi Indicator Cluster Survey – Round 5), SORS and UNICEF. (2013-2014).
16. Economic Impact of Social Enterprises, SORS, EC, Grupa 484, SeCons. (2013-2014).
17. Survey on Corruption and Crime affecting the Business Sector in the Western Balkans, SORS and UNODC. (2012-2014).

18. 2012 – 2015 Survey on Income and Living Conditions - Serbia, SORS, EC and WB. Main responsibilities: Coordinator of the Project team and Chief methodologist: Questionnaire development and testing; Methodology development; Organisation of the fieldwork; Calculation of the indicators. (2012-2015).
19. Prati pare, CRTA. (2012).
20. Citizen budget, under USAID Business Enabling Project in Serbia, BIRN. (2012).
21. Државни динар по становнику, Министарство економије и регионалног развоја. (2012).
22. Business Tendency Survey, SORS and SIDA
23. Fiscal Monitor, Centar za finansije. (2011).
24. MICS4 (Multi Indicator Cluster Survey – Round 4), SORS and UNICEF. (2010-2012).
25. DevInfo Serbia, SORS and UNICEF. (2012).
26. Household survey on experience of corruption and other forms of crime in countries/territories of the Western Balkans, SORS and UNODC. (2010-2011).
27. Inequality and public policy, WIIW. (2008-2010).
28. Innovation Study, World Bank. (2010).

5. ЗАКЉУЧАК И ПРЕДЛОГ

На основу прегледа докторске дисертације под називом „Унапређење званичне статистике применом *Big Data* концепта“, кандидата Тијане Чомић, Комисија констатује да је докторска дисертација написана у складу са свим захтевима стандарда научно-истраживачког рада, као и да испуњава услове предвиђене Законом о високом образовању, стандардима, правилницима и Статутом Факултета организационих наука, Универзитета у Београду.

Тема докторске дисертације је веома актуелна, а резултати добијени истраживањем одговарају предмету и циљевима истраживања. Све хипотезе, постављене у истраживању су тестиране на релевантним примерима из праксе. Такође, резултатима истраживања, потврђена је општа хипотеза, тј. доказано је да се имплементацијом *Big Data* концепта унапређује процес статистичких истраживања, односно да се стварају могућности примене нових метода и техника истраживања, олакшава процес анализе великих количина података и побољшавају укупне перформансе система званичне статистике. Кључни научни допринос дисертације огледају се у унапређењу методолошког поступка процеса статистичког истраживања и формалном опису модела и метода који омогућавају њихову имплементацију.

Истраживање приказано у дисертацији је оригинално, а резултати пружају допринос развоју науке у домену рачунарске статистике. Резултати истраживања могу бити примењени у пракси, у случајевима примене статистичких истраживања. Из приказаног у докторској дисертацији, проистекла су два научна рада, од којих је један

објављен у међународном часопису изузетних вредности, категорије M22. С обзиром на постигнуте резултате, комплексност, актуелност и интердисциплинарност обрађене теме, ова дисертација задовољава највише научне критеријуме и показује способност кандидата Тијане Чомић за самосталан научно-истраживачки рад.

На основу свега наведеног, Комисија предлаже Наставно-научном већу Факултета организационих наука да се докторска дисертација под називом „Унапређење званичне статистике применом *Big Data* концепта“, кандидата Тијане Чомић, прихвати, изложи на увид јавности и упути на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

У Београду, 13.02.2019. године.

ЧЛАНОВИ КОМИСИЈЕ

за оцену завршене докторске дисертације

1. _____
др Зоран Радојичић, редовни професор, Универзитет у
Београду, Факултет организационих наука

Ментор

2. _____
др Александар Ђоковић, доцент,
Универзитет у Београду, Факултет организационих наука

3. _____
др Свјетлана Јанковић Шоја, доцент
Универзитет у Београду, Пољопривредни факултет