

СтановищеВх. № РД-06-322109-05.2018г

за дисертационен труд на Борис Иванов Банков
на тема “Софтуерна система за автоматизирана обработка на неструктурирани данни от социалните мрежи”
за получаване на образователната и научна степен “доктор”
по докторска програма “Информатика”

Изготвил становището: доц. д-р Снежана Сълова, Икономически университет – Варна, доктор по “Приложение на изчислителната техника в икономиката”, доцент по “Информатика и компютърни науки”, член на научното жури по конкурса, съгласно заповед на Ректора на Икономически университет – Варна № РД-06-1177/02.05.2018 г., определена за изготвяне на становище съгласно решение на научното жури от 03.05.2018 г.

Общо представяне на дисертационния труд

Дисертационният труд е с общ обем 188 стр., като включва увод, 3 глави, заключение, съдържащо научните приноси, списък на използваната литература, състоящ се от 105 литературни и 47 интернет източника, 14 приложения и списък на публикациите на автора по дисертацията. Представената дисертация отговаря напълно на изискванията на чл. 27, ал. 2 от Правилника за прилагане на Закона за развитие на академичния състав в Република България.

Преценка на структурата и съдържанието на дисертационния труд

Темата на дисертационния труд е актуална, интересна и значима. В увода е обърнато внимание на нарастването на обема и ролята на неструктурираните данни. Целта, която се поставя в дисертацията, е “да се създаде модел на софтуерна система за обработка на неструктурирани данни, извлечени от социалните мрежи и да се предложи концепция за нейната реализация”. Дефинирани са 4 изследователски задачи и коректно са формулирани тезата, обектът и предметът на научното изследване.

Глава първа е озаглавена “Концепции и модели за обработка на неструктурирани данни в дигитална среда”, има обем 43 стр. В нея се систематизират определения за някои от използваните понятия – данни, информация и знание, извличане на данни, информация и знание, структурирани и неструктурирани данни и т.н. Прави се преглед на методите за обработка на

неструктурирани данни, като се поставя акцент на клъстерния анализ и невронните мрежи. Разглеждат се софтуерните средства за обработка на неструктурирани данни и се достига до извода, че има много нерешени проблеми както в теоретичен, така и в практически аспект и че съществува необходимост от софтуерна система за обработка на текстови масиви от интернет източници, написани на български език.

Глава втора е озаглавена “Модел на софтуерна система за автоматизирана обработка на неструктурирани данни от социалните мрежи ” и е с обем 55 стр. В нея е представена концепция за система за обработка на неструктурирани данни от социалните мрежи. Предвижда се тя да обработва текстове на български език и извличането на данните да става в реално време. Разработен е алгоритъм за определяне на активни клъстери, при клъстеризация на постоянна емисия от данни, чрез въвеждане на променлива, която отчита времето за възстановяване на неактивни клъстери на базата на тяхната значимост.

Глава трета е озаглавена “Софтуерна реализация на системата за автоматизирана обработка на неструктурирани данни от социалните мрежи в Медийна група Черно море” и е с обем 50 стр. Предложената система за обработка на неструктурирани данни от социалните мрежи е апробирана в медийна група „Черно море“ и за социалната мрежа Twitter. Направен е обоснован избор на технологиите за изграждане на системата. Представени са резултатите от направена апробация на системата в периода 01.04 – 16.04.2018 г.

В **заключението** са обобщени постигнатите резултати в изследването.

Дисертационният труд показва, че дисертантът притежава задълбочени теоретични знания по изследваната проблематика. Дисертацията напълно отговаря на изискванията на Закона за развитие на академичния състав в Република България, Правилника за неговото приложение – чл. 27, ал. 1, а също така и на Правилата за условията и реда за придобиване на научни степени и заемане на академични длъжности в Икономически университет – Варна, чл. 34, ал. 2 и 3.

Приложеният **автореферат** отразява точно и вярно съдържанието на дисертационния труд.

Представените публикации са 4, от които 2 самостоятелни статии в индексирани издания и 2 доклада от научни конференции. Те са по темата на дисертацията, отразяват някои от основните идеи на автора. Изискванията за публикации за образователна и научна степен „доктор” по чл. 35, ал. 1, т. 4 на Правилата на ИУ – Варна са спазени.

Идентифициране и оценяване на научните и научно-приложните приноси в дисертационния труд

В представения за рецензиране дисертационен труд могат да се открият следните научни и научно-приложни приноси: предложена е класификация на различните градивни единици на потребителските съобщения за четири от най-популярните социални мрежи; създаден е модел на софтуерна система за автоматизирана обработка на неструктурирани текстови данни на българския език, извлечени от социалните мрежи; разработен е алгоритъм за определяне на активни клъстери, при клъстеризация на постоянна емисия от данни; разработена е концепция за реализиране и е апробирана предлаганата софтуерна система в организация от медийната сфера.

Критични бележки и препоръки

Към дисертационния труд нямам забележки. Препоръчвам на автора да продължи с изследванията си в тази научна област и да публикува резултатите си в авторитетни международни издания.

Въпроси към дисертанта

В какви насоки може да се усъвършенства на предложената система?

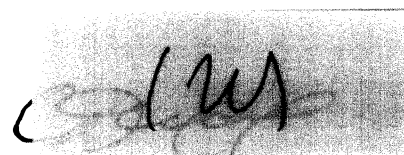
Заклучение

Дисертацията на Борис Банков представлява завършен труд, с който авторът демонстрира способността си за извършване на самостоятелно научно изследване със значими резултати. Разработката има подходящи структурата и съдържание за дисертационен труд и съдържа научни и научно-приложни приноси. Считам, че дисертацията отговаря на изискванията на Закона за развитието на академичния състав в Република България и Правилника за неговото приложение. Давам положителна оценка на дисертацията на тема “Софтуерна система за автоматизирана обработка на неструктурирани данни от социалните мрежи” и предлагам на Борис Иванов Банков да бъде присъдена образователната и научна степен “доктор” по докторска програма “Информатика”.

9.05.2018 г.

гр. Варна

Изготвил становището:



(доц. д-р Сн. Сълова)

СТАНОВИЩЕ

Вх. № РД20-377/21.05.2018г.

от доц. д-р Веселина Господинова Жечева,
Център по информатика и технически науки, Бургаски свободен университет
за дисертационен труд
„СОФТУЕРНА СИСТЕМА ЗА АВТОМАТИЗИРАНА ОБРАБОТКА НА НЕСТРУКТУРИРАНИ ДАННИ ОТ
СОЦИАЛНИТЕ МРЕЖИ”
на докторанта Борис Иванов Банков,
разработен в катедра „Информатика”, Икономически Университет - Варна
по конкурс за присъждане на образователната и научна степен "доктор",
професионално направление 4.6. "Информатика и компютърни науки"

Със Заповед № РД 06-1177 / 02.05.2018 г. на Ректора на Икономически Университет - Варна съм определена за член на научното жури във връзка с конкурс за присъждане на образователната и научна степен "доктор", професионално направление 4.6. Информатика и компютърни науки.

Като член на научното жури получих следните документи:

1. Професионална автобиография;
2. Заповеди за зачисляване в докторантура и отчисляване с право на защита;
3. Протоколи за издържани изпити от докторански минимума;
4. Списък и копия на публикациите за участие в конкурса – 5 броя;
5. Автореферат и дисертационен труд;
6. Декларация за оригиналност на представения дисертационен труд и публикации;
7. Копия на други документи от официалната документация по протичането на докторантурата на ас. Борис Банков.

Данни за дисертационния труд и автореферата. Кандидатът за заемане на образователната и научна степен "доктор" Борис Иванов Банков е представил дисертационен труд, автореферат и четири публикации по темата на дисертацията в научната периодика или в научни сборници.

Представеният дисертационен труд е с обем от 188 страници и се състои от увод, 3 глави, заключение, списък с използвана литература, приложения и списък на публикациите по дисертационния труд. Разпределението на обема на дисертацията е: основен текст – 158 страници, 14 приложения, 52 фигури и 15 таблици.

Тематика на дисертационния труд. Представеният дисертационен труд е посветен на проблем от актуална област, а именно обработката на естествен език, в частност анализ на текст, публикуван в социални мрежи. Работата има научно-приложен характер, като описва методология за решаване на формулирания проблем и разработва практическо средство за решаването му.

В **Увода** са формулирани тезата, целта и задачите на дисертационния труд. Обоснована е необходимостта от обработка на текст на естествен език в Интернет среда с цел откриване на нова и полезна информация. **Целта**, която се поставя в дисертационния труд, е ясна и точна, а именно: да се създаде модел на софтуерна система за обработка на неструктурирани данни, извлечени от социалните мрежи и да се предложи концепция за нейната реализация. **Задачите**, формулирани в увода на дисертационния труд, отговарят на поставената цел и включват анализ на състоянието на теоретичните изследвания и идентифициране на проблеми в областта на автоматизираната обработка на неструктурирани данни, извлечени от социалните мрежи; създаване на модел на софтуерна система за обработка на неструктурирани данни, извлечени от социалните мрежи; разработване на концепция за реализиране и внедряване на системата; апробиране на основните функционалности на предложената система.

Глава 1, озаглавена „Концепции и модели за обработка на неструктурирани данни в дигитална среда“, е обзорна. В нея са разгледани основните понятия и концепции, свързани с обработката на неструктурирани данни, особеностите на различните социални мрежи и техните данни. Описани са технологиите за текстов анализ, като е направен сравнителен анализ на техните предимства и недостатъци, както и достъпните софтуерни средства, които ги реализират. Отбелязано е, че почти липсват средства, които реализират избраната задача на български език.

Глава 2, озаглавена „Модел на софтуерна система за автоматизирана обработка на неструктурирани данни от социалните мрежи“ съдържа описание на модела на предложената система като концептуален модел със състав и структура, основните модули на системата с прилаганите алгоритми, илюстрирани с подходящи примери. Дефинирани са и основните изисквания и функции на системата, като е описано представянето на използваните данни.

В Глава 3, озаглавена „Софтуерна реализация на системата за автоматизирана обработка на неструктурирани данни от социалните мрежи в Медийна група Черно море“ е представена практическата реализация на предложената система. Описани са изискванията на медийна група „Черно море“ към функциите на системата и е обоснован изборът на Twitter като социална мрежа за събиране на тестови данни. Представени са интерфейсът, бизнес логиката и данните на системата и е обоснован изборът на технологични решения за разработка на всеки модул от системата, като са описани прилаганите критерии.

Приноси. Авторът на настоящата рецензия приема дефинираните от докторанта приноси, а именно:

1. Изследвани са източниците, методите, моделите и технологиите за обработка на неструктурираните данни в дигитална среда и е доказана необходимостта от автоматизирана обработка на неструктурирани текстови данни на български език.
2. Анализирани са неструктурираните данни в социалните мрежи и е въведена класификация на различните градивни единици на потребителските съобщения за четири от най-популярните платформи за социално взаимодействие в Интернет.
3. Създаден е модел на софтуерна система за автоматизирана обработка на неструктурирани данни, който е съобразен със спецификата и характера на българския език, използван в публикациите в социалните мрежи.
4. Разработен е алгоритъм за определяне на активни клъстери, при клъстеризация на постоянна емисия от данни, чрез въвеждане на променлива, която отчита времето за възстановяване на неактивни клъстери на база на тяхната значимост.
5. Разработена е концепция за реализиране и частично внедряване на предлаганата софтуерна система в организация от медийната сфера.

Публикации и участия в научни форуми. Представени са четири публикации, от които 2 са статии в списания и 2 – доклади на конференции, като 1 е на английски и 3 на български език. Спазен е необходимият минимум, както и изискването за публикуване в реферирано издание.

Списъкът с използвана литература включва 152 заглавия, от които 105 литературни издания и 47 Интернет ресурси. Библиографията е пълна и актуална, преобладаващата част от източниците са на английски език. Дисертацията е оформена прилежно, всички таблици и фигури са номерирани, стилът е ясен и стегнат. Авторефератът е в обем от 33 страници, като правилно отразява съдържанието на дисертацията.

Представеният дисертационен труд отговаря на изискванията на чл. 34, ал. 2 и ал. 3 на Правилата за условията и реда за придобиване на научни степени и заемане на академични длъжности в Икономически университет – Варна.

Във връзка с дисертационния труд могат да се отправят следните **въпроси, критични забележки и коментари**:


1. Заглавието на подточка 2.4 „Модел на данните“ не съответства точно на съдържанието, което е свързано основно с прилагането на избрани алгоритми за клъстеризация и визуализация на данните.
2. Тъй като системата работи в реално време, как е предвидено да се справя с нарастването на данните? За какъв период от време се предвижда да се обработват? Предвижда ли се разграничаване на постоянно актуални теми и злободневни такива (актуални за кратък период от време)?
3. Полезно за докторанта би било да публикува получените резултати и в чуждестранни реферирани издания, което би увеличило видимостта на работата.

Цялостното впечатление на рецензента от дисертационния труд на кандидата е положително, като приема всички приноси съгласно представената справка. Представената работа показва широка и задълбочена теоретична подготовка и висока степен на изграденост на изследователската теза, яснота относно целите и задачите на изследването, както и професионализъм при разработване на работещ прототип. Представеният труд е актуален и научно подкрепен анализ и съдържа практическо изследване на избраната област. Направените по-горе забележки и коментари нямат за цел да омаловажат неговите постижения, които имат научно-приложен характер и трябва да се възприемат по-скоро като препоръка за бъдещата му работа.

Заклучение

Въз основа на изложеното по-горе считам, че кандидатът Борис Иванов Банков изпълнява всички критерии и изисквания по Закона за развитие на академичния състав в Република България, Правилника за неговото прилагане и Правилата за условията и реда за придобиване на научни степени и заемане на академични длъжности в Икономически университет – Варна. Посочените приноси и положителни страни на дисертационния труд ми дават основание убедено да препоръчам на членовете на уважаемото Научно жури да присъдят на Борис Иванов Банков образователната и научна степен "доктор", професионално направление 4.6 "Информатика и компютърни науки".

Подпис:


/доц. д-р В.Жечева/

16.08.2017 г.

Гр. Бургас

СТАНОВИЩЕ

Вх. № *РД20-380/22.05.2018г.*

от доц. д-р Димитрина Любомирова ПОЛИМИРОВА
Национална лаборатория по компютърна вирусология – БАН

по повод Заповед № 06-1177 / 02.05.2018 г. на проф. д-р Пламен Илиев,
Ректор на Икономически университет – Варна,

определяща състав на научно жури

във връзка с провеждането на защита на дисертационен труд

на Борис Иванов Банков,

**на тема: „Софтуерна система за автоматизирана обработка на
неструктурирани данни от социалните мрежи“**

за придобиване на образователна и научна степен „доктор“ по професионално направление
4.6. – Информатика и компютърни науки, докторска програма – Информатика

с научен ръководител доц. д-р Снежана Сълова

В своето становище ще обърна внимание върху няколко основни неща, чиято оценка е от особено значение при съставянето на определено мнение за качествата на един дисертационен труд. Колегата Борис Банков обръща внимание на развитието в областта на софтуерните технологии и техния потенциал като възможности за автоматизирана обработка на неструктурирани данни от социалните мрежи. Известно е, че българският език е сред езиците с висок риск от отмиране в дигиталната епоха. Конкретният дисертационен труд представя подход за обработка на данни от неструктурирани интернет източници върху текстови масиви на български език.

Темата за *Автоматизирана обработка на неструктурирани данни* като цяло е измежду най-актуалните в съвременните информационни технологии. Тя е свързана с приложения, използващи наличните в Интернет (и по-конкретно в социалните мрежи) динамични ресурси от информация. Разработката има практическа ориентация, като дисертационният труд обхваща научни и научно-приложни резултати, свързани със създаването на реално работещ софтуерен инструмент.

Представените за рецензиране материали са:

- 1) хартиено копие на дисертационния труд, съдържащ 188 страници заедно с приложенията;
- 2) хартиено копие на автореферат от 33 страници;
- 3) хартиени копия на 4 публикации на кандидата, отразяващи резултати в дисертационния труд;
- 4) автобиография на Борис Банков;
- 5) други документи (дипломи, сертификати, заповеди и протоколи), доказващи легитимността на процедурата.

Съгласно ЗРАСРБ, дисертацията трябва да съдържа оригинални приноси, което несъмнено е изпълнено в този случай. Разработеният софтуерен инструмент доказва на практика, че Борис Банков притежава знания и способности за извършване на самостоятелни изследвания, както и за създаване и внедряване на иновативни софтуерни приложения.

Съгласно Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности в ИУ – Варна, кандидатът представя 4 публикации по дисертационния труд, които са публикувани както следва:

- две самостоятелни научни статии: едната в „Известия на Съюза на учените – Варна“, 2017 г., а другата в „Science. Business. Society: International Scientific Journal“, 2018 г.;
- два научни доклада: един самостоятелен в „Предизвикателства пред информационните технологии в контекста на „Хоризонт 2020““, 2016 г., вторият в съавторство в „Икономиката в променящия се свят – национални, регионални и глобални измерения (ИПС – 2017)“, 2017 г.

В научните публикации са публикуване резултатите от дисертацията, които касаят:

- К-средни и клъстеризация на постоянна емисия от данни и алгоритъм за определяне на активни и неактивни клъстери (глава 2);
- избор на технологични решения за разработване на системата и избор на социална мрежа за апробация на системата (глава 3);
- реализиране и внедряване на софтуерната система за обработка на неструктурирани данни от социалните мрежи (глава 3);

Като брой представените публикации удовлетворяват изискванията на Правилника на ИУ – Варна.

Друг важен момент в оценката на представената дисертация е свързан с качествата на нейната **структура**. По мое мнение тя притежава ясна и логична структура, която включва въведение, три глави, заключение, библиография, 14 приложения и публикации по дисертационния труд. Искам да обърна особено внимание на въведението. По начало тази част от всеки докторски труд е сред важните показатели доколко докторантът е наясно с това, което ще изследва и как ще го изследва. Борис Банков се е постарал да включи в своето въведение важни елементи, предхождащи истинското изложение. Дефинирана е ясно тезата, целта на труда и поставените четири задачи, които напълно кореспондират на поставената цел. Дисертационният труд прави добро впечатление със стегнатото си изложение, представено напълно разбираемо и съпроводено с таблици и фигури, чиито брой е добре балансиран.

Друго важно качество на един дисертационен труд е дали авторът му е показал добро **познаване на научната литература и на актуалното състояние на изследвания проблем**. По мое мнение това са важни елементи, които трябва да се търсят в изложението на всяка малка докторска дисертация. Доколкото мога да съдя от изложението на дисертационния труд, Борис Банков е подходил отговорно към натрупаното наследство по въпроса като е демонстрирал необходимата информираност. Цитираната в текста литература обхваща

104 заглавия. Библиографията обхваща 105 литературни източници (от които 16 на кирилица, 89 на английски) и 47 интернет източници.

Разбира се едно от най-важните неща при оценката на даден дисертационен труд са неговите конкретни **приноси**. На първо място бих отбелязала, че едно от несъмнените достойнства на труда е разработването на алгоритъм за определяне на активни клъстери в модул „Клъстеризация“, като Борис Банков въвежда отчитането на променливата „време за възстановяване на неактивни клъстери на база на тяхната значимост“.

Друг важен принос е предложения модел на софтуерна система за автоматизирана обработка на неструктурирани данни от социалните мрежи. По отношение на системата основният принос е в обработката на текстове на български език, използван в публикациите в социалните мрежи.

Научно-приложният принос, който също заслужава внимание да бъде отбелязан, е разработената концепция за реализиране и внедряването на предложената от Борис Банков софтуерна система в организация от медийната сфера. Разработената софтуерна система може да бъде внедрена в други организации с различна дейност и адаптирана към други социални мрежи.

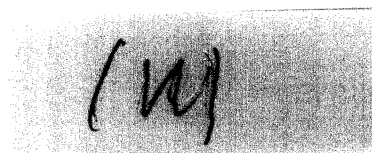
Смятам, че получените резултати доказват наличието на качества за извършване на самостоятелна научна и научно-приложна работа, които са необходимо условие за присъждане на образователната и научна степен „доктор“.

Заключение

Оценката ми за дисертационния труд на Борис Банков е **положителна**. В представения дисертационен труд са получени резултати с научен и научно-приложен характер, които са апробирани. Докторантът демонстрира задълбочени познания в конкретната научна област и научен потенциал за реализация на оригинални идеи.

Постигнатите резултати ми дават основание да предложа на уважаемото научно жури да гласува с положителен вот за присъждане на Борис Банков образователната и научна степен „доктор“.

22.05.2018 г.
гр. София



доц. д-р Димитрина Полимирова