

СТАНОВИЩЕ

от д-р **Маргарита Николаева Терзийска**,
доцент в **Университет по хранителни технологии, град Пловдив**,

на дисертационен труд за присъждане на образователната и научна степен „доктор“
по: област на висше образование 4. Природни науки, математика и информатика
професионално направление 4.6. Информатика и компютърни науки
докторска програма Информатика

Автор: Мария Минева Жекова

Тема: Естественоезиков интерфейс за информационна система от тип „електронен университет“

Научен ръководител: проф. д.м.н. Георги Атанасов Тотков

1. Общо представяне на процедурата и докторанта

Утвърдена съм със заповед № РД-Д1-616/25.03.2022 г. на проф. д-р Румен Младенов, ректор на ПУ „Паисий Хилендарски“, за член на Научно жури във връзка с процедура за придобиване на образователна и научна степен „доктор“ по: област на висше образование 4. Природни науки, математика и информатика, професионално направление 4.6 Информатика и компютърни науки, докторска програма Информатика от **Мария Минева Жекова** с дисертация на тема *Естественоезиков интерфейс за информационна система от тип „електронен университет“*, научен ръководител проф. д.м.н. Георги Атанасов Тотков.

Като член на Научното жури съм получила:

- молба до Ректора на ПУ за разкриване на процедурата за защита на дисертационен труд;
- автобиография в европейски формат;
- протокол от катедрения съвет, свързан с докладване на готовността за откриване на процедурата и с предварително обсъждане на дисертационния труд;
- дисертационен труд;
- автореферат;
- списък на научните публикации по темата на дисертацията;
- копия на научните публикации;
- декларация за оригиналност и достоверност на приложените документи;
- справка за спазване на специфичните изисквания на ФМИ.

При оценяването на дисертационния труд под внимание следва да се вземат изискванията на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за неговото прилагане (ППЗРАСРБ) и Правилника за развитие на академичния състав на ПУ(ПРАСПУ) и специфичните изисквания на ФМИ. Основните норми, които трябва да бъдат спазени са:

1. Съгласно чл. 6(3) от ЗРАСРБ „дисертационният труд трябва да съдържа научни или научно приложни резултати, които представляват оригинален принос в науката. Дисертаци-

онният труд трябва да показва, че кандидатът притежава задълбочени теоретични знания по съответната специалност и способности за самостоятелни научни изследвания“.

2. Според чл. 27(2) от ППЗРАСРБ дисертационният труд трябва да се представи във вид и обем, съответстващи на специфичните изисквания на първичното звено. Дисертационният труд трябва да съдържа: заглавна страница; съдържание; увод; изложение; заключение – резюме на получените резултати; библиография.

Дисертацията напълно покрива тези изисквания като се състои от заглавна страница, съдържание, увод, изложение под формата на литературен обзор и три глави с приносите на дисертанта и заключение.

3. Съгласно чл. 36(1) от ПРАСПУ, специфичните изисквания на ФМИ при ПУ са кандидатът да има:

- минимум 4 публикации в рецензирани издания;
- минимум 1 публикация в списание.

Докторантът е представил 6 публикации, свързани с дисертационния труд, като всичките са в рецензирани издания, а три от тях са в списания.

2. Актуалност на тематиката

Компютърната обработка на естествен език (NLP) е едно от най-предизвикателните и бързоразвиващи се направления от изкуствения интелект (AI). Благодарение на него в ежедневието си използваме чатботове, гласови асистенти, преводачи и редица други инструменти. Все повече компании виждат ползите и използват NLP за извличане на информация от големи обеми данни и за автоматизиране на повтарящи се дейности. С нарастването на неструктурираните данни задачите пред NLP се увеличават като стремежът е да се разберат точно нюансите, контекста и неяснотите в човешкия език и в крайна сметка NLP да се превърне в технология, достъпна за всички, дори и за неспециалисти. Именно това, е и основната цел на дисертационното изследване, което свидетелства за неговата актуалност и значимост.

3. Познаване на проблема

В дисертационният труд са цитирани общо 82 източника, като 8 от тях са на кирилица и 74 са на английски език. Прави впечатление, че по-голямата част от препратките са към по-стари, но важни за тематиката източници. Около 45% от цитираните източници са от последните 10 години. Въпреки това, от начина, по който е представен обзорът в първа глава на дисертационния труд, личи, че докторантът се е запознал в детайли с научната проблематика.

4. Методика на изследването

Избраната от докторанта методика за провеждане на изследването съответства на поставената цел и произтичащите от нея задачи. Дизайнът на ЕЕИ е реализиран като ASP.NET приложение, като за целта са използвани възможностите на езика C# и .NET framework. За работа с БД са използвани съответно Microsoft SQL Server и SQL Server Management Studio. Процесът на преобразуване на потребителски текст на естествен език в заявка към БД е реализиран на Python. Езикът е избран, тъй като поддържа голям брой библиотеки, свързани с об-

работката на лингвистични данни. Част от тях са използвани при разработката на ЕЕИ, като напр. *punkt*, *request*, *nltk*, *google_trans_new*, *re* и т.н. Всичко това показва, че избраната методика е напълно адекватна на целта на работата.

5. Характеристика и оценка на дисертационния труд и приносите

Дисертацията се състои от четири глави (включващи увод и заключение), списък на цитираната литература, списък на публикациите свързани с дисертацията, списък на приносите за които претендира дисертантът. Като приложения са дадени лингвистичен модел за прилагането на ЕЕИ към прототип на БД на НАОА (експеримент 2) и резултати от проведените тестове.

Глава първа представлява увод в предметната област на изследването. Въведени са някои основни понятия, направен е задълбочен анализ на съществуващи инструменти и системи с ЕЕИ за естественоезикова обработка, като е обърнато внимание на подходите за моделиране и структуриране на информационния контекст и на методите за програмната им обработка. Формулирани са целта и задачите на дисертационния труд.

Във втора глава е предложен модел на процес на създаване на ЕЕИ към ИС и е представена неговата архитектура, методи и модели. Предложена е също така методика за моделиране на предметна област, свързана с ИС.

В глава 3 е описана софтуерната реализация на ЕЕИ.

В глава 4 са представени експерименталните резултати от прилагането на ЕЕИ към две различни ИС в областта на висшето образование – ЕЕИ към БД на еУниверситет и ЕЕИ към прототип на БД на НАОА.

Приемам приносите така, както са представени в дисертационния труд:

Научни приноси:

1. Предложена е методика за създаване на ЕЕИ към информационна система.

Научно-приложни приноси:

1. Разработена е архитектура на софтуерен инструмент с ЕЕИ към ИС;
2. Реализиран е софтуерен инструмент с ЕЕИ към ИС в област ВО с различно предназначение и обхват.

Приложни приноси:

1. Създадени и тествани са компютърни модели на ПО по предложената методика;
2. Създадени са експериментални ЕЕИ за две ИС в област ВО, в които са проведени тестове с потребителски въпроси.

6. Преценка на публикациите и личния принос на докторанта

Във връзка с дисертацията, кандидатът е представил 6 публикации. Публикация 4 е в издание, индексирано в Scopus, а публикация 6 е в издание, индексирано в Web of Science. Дисертантът е участвал в 6 конференции, където е докладвал получените по дисертацията резултати. В представените материали няма данни за цитирания. Две от публикациите по дисертацията са самостоятелни, а в останалите докторантът е първи съавтор. Това ми дава основание да смятам, че неговото участие е съществено.

7. Автореферат

Представеният автореферат е в обем от 36 страници и е структуриран в съответствие с дисертационния труд.

8. Препоръки за бъдещо използване на дисертационните приноси и резултати

Дисертацията е много добре представена. Написана е стегнато и ясно, изложението е логически последователно. Въпреки това, имам следните въпроси, забележки и препоръки:

1. Защо за създаване на ЕЕИ са използвани различни езици за програмиране, а не само един, например Python?

2. В Заключението на стр. 134 се твърди, че експерименталните ЕЕИ "се справят със задачата по извличане на информация и получаване на отговор от БД през ЕЕИ", както и "създаденият софтуерен модул е достъпен, адекватен и устойчив", но в дисертационния труд няма доказателство за тези твърдения. Направено ли е сравнение за получените резултати, върнати след заявка от ЕЕИ и например от стандартна заявка към БД?

3. Липсва демонстрация на работата на системата. Добре би било да има поне няколко изгледа (скриншота) от типа „въведен въпрос-отговор“. Това би могло да стане след фиг. 15 на стр. 65, на която е представен потребителския интерфейс на ЕЕИ или в Приложение 2, където е представен резултат от експеримента до заявка към базата данни, но липсва отговора т.е. какво вижда като отговор клиента.

4. Фиг. 14 на стр. 58 представя "Архитектура на ЕЕИ към информационна система". Дали е коректно представено, че ЕЕИ и ИС се намират при клиента? Не трябва ли да са на уеб сървъра?

5. Препоръчвам на докторанта да продължи да работи по тази изключително актуална тематика като се опре на по-нови и съвременни разработки и обърне съществено внимание на дълбоките невронни мрежи, които към момента дават най-добри резултати при обработка на естествен език.

ЗАКЛЮЧЕНИЕ

Представеният дисертационен труд отговаря напълно на съвкупността от критерии и показатели за придобиване на образователна и научна степен "Доктор", съгласно Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за неговото прилагане и Правилниците за условията и реда за придобиване на научни степени и заемане на академични длъжности на ПУ "Паисий Хилендарски" и Факултета по математика и информатика. Личното ми мнение е, че дисертацията е разработена на много добро ниво. Разглежданият в нея проблем е актуален. Приносите са значими. Като имам предвид всичко това, **оценявам положително** дисертационния труд и предлагам на научното жури да присъди образователната и научна степен "доктор" на **Мария Минева Жекова** по професионално направление 4.6 Информатика и компютърни науки, докторска програма "Информатика".

4.05.2022 г.

Изготвил становището:

/доц. д-р инж. М. Терзийска/