

# Рецензия

от

проф. д-р Стефка Стоянова Фиданова

Институт по Информационни и Комуникационни Технологии

Българска Академия на Науките

направление 4.6 „Информатика и компютърни науки”

за

придобиване на научната степен „доктор на науките“

по професионално направление 2.1 Филология

с кандидат гл. ас. д-р Велислава Райкова Стойкова

Утвърдена съм със заповед No 3 – РК- 268/13.07.2022 на ректора на Нов Български Университет за член на научно жури за придобиване на научната степен „доктор на науките“ по професионално направление 2.1 Филология от д-р Велислава Райкова Стойкова с дисертация на тема „Machine Translation Approaches to the Bulgarian Language“ (Подходи за машинен превод на български език).

Като член на Научното жури съм получила:

1. Дисертация за придобиване на научна степен „доктор на науките“ по направление 2.1 „Филология“;
2. Автореферат;
3. Професионална биография;
4. Списък с публикациите на автора, свързани с дисертацията.

При оценяването на дисертационния труд се вземат под внимание изискванията на Закона за развитие на академичния състав в Република България (ЗРАСРБ), Правилника за неговото прилагане (ППЗРАСРБ), съответстващия Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности в Нов Български Университет и Наредба за развитието на академичния състав на НБУ.

- Съгласно чл. 6 (3) от ЗРАСРБ „дисертационният труд трябва да съдържа научни или научно-приложни резултати, които представляват оригинален принос в науката. Дисертационният труд трябва да показва, че кандидатът притежава задълбочени теоретични знания по съответната специалност и способности за самостоятелни научни изследвания“.
- 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 се състои от увод, 3 глави, заключение, библиография, приложения и справочен апарат (списъци с използвани съкращения, индекс на основните термини и др.). Представени са списъци на авторските публикации по дисертационния труд, списък на забелязаните цитирания и на проектите в които кандидатката е участвала.

#### **1. Актуалност на проблема и целесъобразност на поставените цели и задачи**

Живеем във времена, когато компютрите влияят на всички сфери на нашия живот. Бурното развитие на много области, през последните десетилетия, се дължи на навлизане на компютърното подпомагане. Това води до нуждата от бърз и качествен превод. Развитието на езиковите технологии подпомага разработването на компютърни програми за машинен превод. Разработването на такива технологии започва, преди няколко десетилетия, с разработки за английски език. За съжаление изследванията в тази област, свързани с български език са все още в началото си. Това прави предложения дисертационен труд особено актуален.

Представеният ми за рецензиране дисертационен труд е насочен основно към анализ на съществуващите подходи в областта на езиковите технологии и възможностите и ефективността им за използването им за български език.

Основната цел на дисертационния труд е изследване и анализ на възможностите на формалните езикови технологии за компютърен анализ на български език, както и възможностите им за използване за машинен превод.

За постигането на поставената цел са дефинирани следните задачи:

- Изследване на разнообразни подходи за машинен превод на български език с използването на контрастивен анализ.
- Сравняване на компютърни програми за български език и резултатите от работата им, отчитайки представянето на езиковите фрагменти; времето и ресурсите, които използват; качеството на превода и възможностите за усъвършенстване и развитие.

## **2. Познание състоянието на проблема от страна на дисертантката**

Няма съмнение, че дисертантката е навлязла много добре в научната проблематика. 30% от цитираните литературни източници са от последните 10 години, което прави този труд много актуален. От друга страна в него присъстват и позовавания на по-стари източници, които са основни за научната проблематиката. Общият брой цитирани източници е 107, като 89 са на английски език, 16 са на български и 2 на руски.

Познаването на проблема от страна на дисертантката и нейният изследователски опит са много добре илюстрирани от глава 1 на дисертационния труд. Там са разгледани общите положения при машинния превод.

## **3. Методика на изследването**

Методиката за провеждане на изследването, избрана от дисертантката, произтича от поставената цел и съответства на произтичащите от целта задачи. Тя е напълно адекватна за темата на работата. Авторката използва език за представяне на лексикална информация DATR, Универсален мрежови език и Sketch Engine. Тези подходи са били разработвани основно за английски език. Авторката разглежда техните възможности и необходимите изменения и доразвиване, за да могат да бъдат успешно прилагани и за български език.

## **4. Характеристика и оценка на приносите на дисертационния труд**

Дисертацията започва с увод в предметната област на изследването. Там е дадена мотивация за избор на проблема и обекта на изследването. Направен е кратък преглед на използваните в дисертацията методи. Посочени са целите и задачите на

дисертацията. Приложени са таблици и фигури, илюстриращи разглежданите задачи и постигнатите резултати.

Глава 2 представя формалния език DATR и неговите възможности за прилагането му за машинен превод на български език и нужното доразвиване, за да бъде преводът ефективен.

Глава 3 е посветена на Универсалния мрежов език и използването му за машинен превод. Той е използван за представянето на езикови явления в български език. Направено е сравнение с английски и руски език, като е демонстриран механизма на машинния превод между тях. Представени са и примери за машинен превод между български и словашки език, използващи Универсалния мрежов език.

В Глава 4 е разгледан метод за търсене в електронни двуезични бази данни с използване на статистически подход. Статистическите подходи дават добри резултати при търсене в големи бази данни и по-специално успоредни и сравними двуезични електронни корпуси. Те са бързи и в същото време дават добро качество на получените резултати. Използваният подход е подкрепен с примери за машинен превод на български, словашки и сръбски език.

Дисертантката е отбелязала следните приноси:

1. Даден е анализ на подходите за машинен превод на български език въз основа на използването на формалните езикови теории DATR и Универсалния мрежов език, съответните им програмни приложения и въз основа на статистическите подходи на Sketch Engine.

2. За основа на машинния превод на български език е прието съответствието между семантичен еквивалент, логически еквивалент и преводен еквивалент.

3. За анализа на програмите за машинен превод на български език са ползвани както изходните, така и получените езикови данни, които са подкрепени с убедителни примери.

4. Даден е подход за компютърно измерване на коректността на машинния превод на български език, както и за подобряване на преводния резултат.

Дадени са насоки за възможно бъдещо развитие на тематиката.

## **5. Значимост на разработката за науката и практиката**

Извършената от дисертантката работа е достатъчна по обем и задълбоченост на изследването. Показаните подходи са с практическа реализация. В този смисъл намирам работата за значима както в научно, така и в практическо отношение.

## **6. Преценка на публикациите по дисертацията**

Във връзка с дисертацията, кандидатката е представила 12 публикации. 7 от тях са в издания реферирани и индексирани в световната система за реферирани и индексирани, като 4 са самостоятелни. 5 са в рецензирани нереперирани издания, като 4 от тях са самостоятелни. Общият брой точки по група показатели Г е 200 при изискуеми 100. По показатели от група Д Велислава Стойкова е представила 7 цитирания, в реферирани издания на нейни публикации, с което общият брой точки е 105 при изискуеми 100. Всички публикации и цитирания са след като кандидатката е защитила образователната и научна степен „доктор“. Това е доказателство, че те не са използвани в предходни процедури.

Според мен, представената ми дисертация отговаря на изискванията на ЗРАСРБ, както и на правилника за неговото прилагане като представя авторско изследване по значим научен проблем – машинния превод, което е със съществени научни и приложни резултати. Авторът на дисертацията отговаря на и дори надхвърля наукометричните показатели по отношение на публикациите и цитиранията, представени към дисертацията.

## **7. Лично участие на кандидата**

8 от представените по дисертацията публикации са самостоятелни, което по категоричен начин показва личното участие на кандидатката.

## **8. Автореферат**

Като цяло, авторефератът правилно отразява съдържанието на дисертацията.

## **9. Критични бележки**

Задачите на дисертацията не са строго формулирани. Въпреки, че авторефератът правилно отразява съдържанието на дисертацията и приносите ѝ, той не следва точно структурата на дисертацията.

## 10. Лични впечатления

Познавам отдавна Велислава Стойкова. Тя е един задълбочен учен, който умело съчетава познанията си в областта на българския език и новите компютърни технологии. Тя е един от първите в България, които въведоха използването на компютърни технологии за анализ и обработка на текстове. Велислава Стойкова е участвала успешно, както се вижда от публикациите ѝ, в конференции в областта на филологията и в конференции в областта на приложната математика и информатика. Това показва колко плодотворно може да бъде комбинирането на тези научни направления.

## 11. Заключение

Като следствие на изложеното по-горе, може да се констатира, че са изпълнени всички изисквания на Закона за развитие на академичния състав (ЗРАСРБ), Правилника за неговото прилагане (ППЗРАСРБ), съответстващия Правилник за условията и реда за придобиване на научни степени и заемане на академични длъжности в НБУ и Наредбата за развитието на академичния състав на НБУ. Мога да заявя, че нивото на тази дисертация удовлетворява и значително надхвърля минималните изискванията.

Забелязаните от мен неточности и направените забележки не намаляват значимостта на получените резултати и научната стойност на предоставения ми труд.

Всичко това ми дава основание за положителна оценка и предлагам на почитаемото Научно жури да присъди научната степен „доктор на науките“ по професионално направление 2.1 „Филология“ на Велислава Райкова Стойкова.

25.07.2022 год.

гр. София

(проф. д-р Стефка Фиданова)