

СТАНОВИЩЕ

относно

дисертационния труд на **Александър Николаев Попов** за придобиване на научната степен „доктор” по докторска програма „Информатика” в професионално направление 4.6 „Информатика и компютърни науки” на тема „**Моделиране на лексикалното знание с цел автоматична обработка на естествен език**”

от

доц. д-р Геннадий Павлович Агре,

Институт по информационни и комуникационни технологии – БАН

Със заповед №127/12.07.2018г. на Директора на Института по информационни и комуникационни технологии – БАН съм утвърден за член на Научното жури във връзка с процедура за придобиване на образователна и научна степен „доктор” по докторска програма „Информатика” в професионално направление 4.6 „Информатика и компютърни науки” от Александър Николаев Попов с дисертация на тема „Моделиране на лексикалното знание с цел автоматична обработка на естествен език” с научен ръководител доц. д-р Кирил Симов.

Като член на журито получих:

1. Дисертация за придобиване на образователна и научна степен „доктор” в професионално направление 4.6 „Информатика и компютърни науки” (на английски език)
2. Автореферат (на български език)
3. Професионална биография
4. Копия от публикации на автора, свързани с дисертацията.

Дисертацията съдържа 145 страници, структурирани в Увод – глава първа, 7 глави, заключение, библиография и три приложения.

1. Актуалност на тематиката на дисертационния труд и целесъобразност на поставените цели и задачи

Дисертацията е свързана с моделиране на лексикалното знание по отношение на естествени езици – задача, стояща в основата на решаване на такива централни за областта на обработка на естествени езици проблеми като машинен превод, снемане на лексикална многозначност, изчисляване на сходството на думи и др. От тази гледна точка тематиката на дисертацията е напълно актуална, а поставените за решаване цели и задачи са целесъобразни.

2. Познание състоянието на тематичната област от страна на дисертанта

Дисертантът показва много добро познание на проблемната област както и на тематики, свързани с реализацията на целите на дисертацията. В третата глава е направен аналитичен обзор на основните подходи за представяне на лексикалното знание с цел автоматична обработка. Библиографията съдържа 131 източника, като трябва да се отбележи, че дисертантът е автор на обширен обзор за използване на рекурентни мрежи за снемане на лексикалната многозначност, публикуван в списание с SJR ранг на SCOPUS.

3. Методика на изследването

Избраната от дисертанта методика за провеждане на изследването произтича от поставените в дисертацията цели и задачи. Убедително е доказана адекватността на използване на релационни знания за моделиране на лексикона, както и използване на архитектури, базирани върху рекурентни невронни мрежи, за решаване на задачата за снемане на лексикалната многозначност.

4. Кратка аналитична характеристика на дисертационния труд

Първата глава представлява увод в тематичната област на изследването. В нея е дадена мотивацията за избор на тематиката и са посочени целта и задачите на дисертацията. Във втората глава са дадени по-формални дефиниции на поставените задачи: определяне на частите на речта, снемане на лексикалната многозначност и определяне на сходство и свързаност между думи. Третата глава анализира основните резултати, постигнати в избраната предметна област. В началото са детайлно разгледани явните модели за представяне на лексикални знания. След това е направен аналитичен обзор на съществуващите методи за снемане на лексикалната многозначност и на използването на невронни мрежи за решаване на различни задачи, свързани с обработка на естествен език. Основните изследвания по дисертацията са представени в следващите четири глави. Глава 4 е посветена на предложеното от дисертанта решение на задачата за определяне на части на речта. Глава 5 представя работата по обогатяване на структурата на популярния семантичен ресурс WordNet, както и оценка на новите ресурси от гледна точка на тяхното използване от базираните на знания методи за снемане на лексикалната многозначност. Глава 6 разширява представената в предишната глава работа в посока научаване на разпределени модели на думи и техния смисъл. Глава 7 представя два алтернативни подхода за решаване на задачата за снемане на лексикалната многозначност чрез използване на рекурентни невронни мрежи. Глава 8 анализира възможността за използване на хибридни архитектури, базирани на успоредно обучение (multi-task learning), за решаване на различни задачи, свързани с моделиране на лексикалното знание. Заключението обобщава съдържанието и резултатите на дисертацията, представя приносите на дисертанта, неговата публикационна дейност, информацията за участие на дисертанта в различни проекти и конференции за представяне на резултатите, както и бъдещи насоки на работата.

5. Приноси на дисертационния труд

Основните научни и научни-приложни приноси са посочени коректно. Резултатите са получени и използвани в рамките на два международни проекта по 7-та рамкова програма и на един национален проект с Фонда за научни изследвания, което представлява една допълнителна гаранция за тяхното качество.

6. Преценка на публикациите по дисертационния труд

Дисертантът представи 14 публикации по темата на дисертацията, като всички те са публикувани в реферирани издания: 8 публикации са реферирани в SCOPUS (5 от тях – във Web of Science), а останалите 6 – в трудове на международни конференции и семинари. Публикациите отразяват коректно съдържанието на дисертацията и представляват оригинални постижения на дисертанта.

Съгласно действащият в момента Правилник за специфичните условия за придобиване на научни степени и за заемане на академични длъжности в ИИКТ-БАН, дисертантът трябва да има най-малко 3 публикации по темата на дисертацията, от които най-малко една да е в списание с импакт фактор/ранг или в специализирано международно издание.

Съгласно Правилника за прилагане на Закона за развитието на академичния състав в Република България, който е в сила от 06.07.2018 г., минималните национални изисквания за придобиване на научна степен „доктор” е 30 точки, събрани от публикации в издания, които са реферирани и индексирани в световноизвестни бази данни с научна информация, и в нереферирани списания с научно рецензиране или в редактирани колективни токове. Съгласно приведените в Правилника формули, публикациите на дисертанта, индексирани в SCOPUS, набират 156,65 точки, а останалите публикации – 59 точки.

Това показва, че публикационната дейност на дисертанта значително надхвърля както минимални национални изисквания, така и специфичните изисквания на Института по информационни и комуникационни технологии – БАН за придобиване на научна степен „доктор”.

Освен това, трябва да се отбележи, че 5 от публикациите на дисертанта имат общо 8 забелязани цитирания от чуждестранни автори.

7. Лично участие на докторанта

Личното участие на докторанта е значително – за това говори фактът, че докторантът има 4 самостоятелни публикации (2 от тях са индексирани в SCOPUS и Web of Science), в 1 публикация той е първият автор, в 5 публикации – на второ място и в 4 – на трето.

8. Автореферат

Авторефератът отразява коректно целите, задачите и резултатите от дисертационния труд.

9. Заключение

От всичко приведеното по-горе следва, че са напълно изпълнени всички изисквания на Закона за развитие на академичния състав (ЗРАСРБЧ), Правилника за неговото прилагане (ППЗ) и Правилника за специфичните условия за придобиване на научни степени и за заемане на академични длъжности в ИИКТ-БАН. Смятам, че дисертацията на Александър Николаев Попов има всички качества, изисквани от един дисертационен труд. Получените от дисертанта резултати са значителни и допринасят за развитие на съществуващите методи за моделиране на лексикалното знание.

Всичко това ми дава основание за положителната оценка и убедено препоръчвам на почитаемото Научно жури да присъди образователната и научна степен „доктор” в професионалното направление 4.6 „Информатика и компютърни науки” на Александър Николаев Попов.

10 октомври, 2018 г.
София