

СТАНОВИЩЕ

на проф. д-р Мария Нишева – ФМИ, СУ „Св. Климент Охридски“

за дисертационен труд за придобиване на образователната и научна степен
“доктор” по професионално направление 4.6 „Информатика и компютърни
науки“

на тема „Моделиране на лексикалното знание с цел автоматична обработка на
естествен език“

с автор Александър Николаев Попов

Със заповед № 127/12.07.2018 г. на Директора на Института по информационни и комуникационни технологии при БАН съм назначена за член на научно жури във връзка с процедурата за придобиване на ОНС „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“, докторска програма „Информатика“, от Александър Николаев Попов с дисертационен труд на тема „Моделиране на лексикалното знание с цел автоматична обработка на естествен език“.

1. Обща информация за докторанта

Александър Попов има висше образование – ОКС „бакалавър“ по английска филология, завършена в СУ „Св. Климент Охридски“, и ОКС „магистър“ по компютърни науки, завършена в Imperial College, Лондон.

Работил е и продължава да работи като програмист и разработчик на софтуер за обработка на естествен език. Има преподавателски опит и опит в изпълнение на международни научноизследователски проекти.

Научните интереси и досегашните изследвателски резултати на Александър Попов са в областите извличане на знания, обработка на естествен език, семантични технологии.

2. Оценка за съответствие с минималните национални изисквания и с допълнителните изисквания по чл. 1а, ал. 2 от ППЗРАСРБ

Минималните национални изисквания за получаване на ОНС „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“ според Приложението към чл. 1а, ал. 1 от ППЗРАСРБ включват натрупване на поне 30 точки по Група от показатели Г.

Съгласно Правилника за специфичните условия за придобиване на научни степени и за заемане на академични длъжности в Института по информационни и комуникационни технологии, дисертацията за получаване на ОНС „доктор“ трябва да е базирана на поне три научни публикации, поне една от които да е в списание с импакт фактор или в специализирано международно издание.

Резултатите от дисертационния труд на Александър Попов са представени в 14 публикации, от които:

- осем публикации са в специализирани международни издания, реферирани и индексирани от Scopus (като пет от тях са в издания, реферирани и индексирани и от Web of Science);
- шест публикации са в сборници с трудове на международни научни събития, които не се реферират от Scopus или Web of Science.

Забелязани са общо осем цитирания на трудове на автора по резултати от неговата дисертация – три в издания, реферирани и индексирани от Scopus, и пет в колективни томове с научно рецензиране.

Така кандидатът натрупва общо 215,65 точки по Група от показатели Г (156,65 точки от публикации в реферирани издания плюс 59 точки от публикации в нереферирани издания) и 30 точки по Група от показатели Д.

Следователно, минималните национални изисквания и допълнителните изисквания по чл. 1а, ал. 2 от ППЗРАСРБ за получаване на ОНС „доктор“ по професионално направление 4.6 „Информатика и компютърни науки“ са покрити напълно и надхвърлени от кандидата.

3. Анализ на съдържанието, резултатите и приносите на дисертационния труд

Дисертационният труд е посветен на изследвания в областта на обработката на естествен език. Негова основна цел е анализирането на различни аспекти на лексикалното знание и разработването на дистрибутивни модели на лексикално знание, които обхващат и знание, извлечено от теорията.

Дисертацията се състои от 150 страници текст на английски език, включващ девет глави, списък на използвани литературни източници и три приложения.

В уводната първа глава е представена областта на изследванията и са формулирани и обосновани целите на дисертационния труд.

Втора глава дефинира на подходящо ниво на абстракция основните задачи, които произтичат от целите на дисертацията.

Трета глава има обзорен характер и съдържа преглед на достигнатите основни резултати от изследванията в избраната област.

Четвърта глава е посветена на изследване на рекурентните невронни мрежи във връзка с тяхната приложимост като инструмент за анализ на последователности от

словоформи. Предложени са дистрибутивни представяния за български език и модел за представяне на морфологична информация, която има ключово значение за коректно решаване на задачата за определяне на частите на речта.

В пета глава са представени резултати от осъщественото от автора изследване на хипотезата, че обогатяването на т. нар. графи със знание с допълнителни парадигматични и синтагматични релации може да се извърши чрез извличане на информация от текстови, полуструктуриран и структурирани ресурси.

Шеста глава дискутира получени резултати във връзка с построяването на дистрибутивни представяния на различни видове лексикални единици.

В седма глава са представени резултати на автора от изследване на приложимостта на рекурентни невронни мрежи за снемане на лексикалната многозначност.

Осма глава е посветена на резултати от успоредно обучение на рекурентни невронни мрежи върху различни задачи за обработка на естествен език, основаващи се на лексикално знание.

В заключителната девета глава е направено обобщение на получените резултати от изследванията на автора по темата на дисертационния труд.

Основните *научни и научно-приложни приноси* на дисертационния труд могат да бъдат формулирани както следва:

- Предложен е подход за моделиране на лексикона чрез релационни знания и е показана валидността на този подход;
- Показано е, че кодирането на релационна лексикална информация във вероятностни модели е подходящ инструмент за лексикално моделиране;
- Предложени са оригинални методи за представяне на лексикона чрез векторни модели, както и иновативен подход за представяне на морфологична информация;
- Предложен е нов подход за снемане на лексикалната многозначност с обучаващ сигнал;
- Предложени са архитектури на рекурентни невронни мрежи за трансформиране на поредици от данни;
- Генерирани са множества от нови релации между значения на думи във формата на WordNet;
- Обучени са и са оценени множество модели за векторно представяне.

Дисертационният труд прави отлично впечатление с обхвата, задълбочеността и аргументираността на изложението. Областта на изследванията е актуална и сложна и постигането на съществени резултати в нея изисква много добра интердисциплинарна подготовка, постоянна и интензивна работа. Постигнатите резултати са оригинални и значими и съответстват на поставените цели.

4. Публикации, които отразяват дисертацията. Отражение на резултатите на дисертацията в трудове на други автори

Резултатите, получени в дисертационния труд, са отразени в 14 научни публикации, сведения за които се съдържат в т. 2 на настоящото становище. Четири от публикациите са самостоятелни, а останалите 10 са в съавторство. Нямам никакво съмнение в същественния личен принос на докторанта в колективните публикации.

Посочени са данни за осем забелязани цитирания на трудове на автора по резултати от дисертационния му труд. Междувременно беше публикувана в реферирано международно издание още една статия на други автори, която цитира две публикации на Ал. Попов, отразяващи части от неговата дисертация.

5. Критични бележки и препоръки

Нямам съществени забележки към съдържанието на дисертацията. Като недостатък на дисертационния труд бих посочила липсата на кратки обобщени формулировки на неговите приноси.

6. Автореферат

Авторефератът отразява пълно и точно съдържанието и приносите на дисертационния труд, но обемът му е твърде голям. Вероятно това се обяснява с решението на автора да напише дисертацията си на английски език и стремежа да представи и достатъчно детайлизиран автореферат на български език.

7. Заключение

Въз основа на изложеното по-горе смятам, че дисертационният труд напълно удовлетворява и надхвърля изискванията на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ и Правилника за специфичните условия за придобиване на научни степени и за заемане на научни длъжности в ИИКТ при БАН. Докторантът е постигнал значими изследователски резултати, които представляват оригинален принос в избраната научна област. **Убедено предлагам на уважаемото научно жури да присъди на автора му Александър Николаев Попов образователната и научна степен „доктор” в професионално направление 4.6 „Информатика и компютърни науки“.**

София, 08.10.2018 г.

