

## **СТАНОВИЩЕ**

за дисертацията на **Ивелина Мирчева Николова**

на тема

**„Приложение на обработката на естествен език за изграждане на семантични системи”,**

представена за присъждане на образователната и научна степен **“доктор”**

по научна специалност

**01.01.12 „Информатика“**

в професионално направление

**4.6 „Информатика и компютърни науки“**

от проф. д-р **Радослав Димов Павлов,**

**Институт по математика и информатика - БАН**

Това становище е написано и представено на основание на заповед на Директора на ИИКТ № 168 от 08.10.2014, както и на решението на научното жури по процедурата. То е изготвено въз основа на ЗРАСРБ, Правилника за прилагане на ЗРАСРБ, Правилника за условията и реда за придобиване на научни степени и за заемане на академични длъжности във ИИКТ и указания за изготвяне на рецензии и становища от членове на научни журита и за

документите в електронен вид, подавани от кандидата по процедури за академични длъжности и научни степени на ИИКТ.

Представеният ми комплект от материали на книжен и на електронен носител включва всички предвидени документи в посочения по-горе закон и неговите приложения.

Темата на дисертационния труд на Ивелина Мирчева Николова е в една актуална и бързо развиваща се област на научни изследвания за обработка на естествен език и изграждане на семантични системи с приложения при автоматичната обработка на клинични биомедицински текстове. Обект на научно и научно-приложно изследване са средства за автоматично разпознаване на парафрази в компютърната лингвистика и разработването на съответен прототип за биомедицински записи; средства за автоматично извличане на релации между понятия от текст и разработването на декларативен концептуален модел, както и на съответен прототип за медицински приложения. Разработените прототипи са интегрирани в семантични системи и са експериментирани подходи, базирани на правила и на машинно самообучение.

Дисертационният труд на Ивелина Мирчева Николова е в обем 129 страници и се състои от общо увод, 4 глави, заключение, списък на използваните термини, съкращенията и означения, списъци на използваните фигури и таблици и библиография.

В Увода е разгледана актуалността на проблематиката и са представени целите и задачите, поставени в дисертацията.

В Глава 1 е представен преглед на съществуващи методи и програмни продукти в областта. Разгледани са съществени характеристики на семантичните системи, базирани на онтологии, проблемите при автоматичното разпознаване на парафрази, автоматичното извличане на понятия и релации между тях и в частност при използването на UMLS, както и структурирането на описания на електронни здравни записи.

В Глава 2 е обоснована необходимостта от създаване на подходящ концептуален модел, в който се дефинират основни понятия от предметната област и релациите между тях. Представени са методи за разпознаване и извличане на синоними и парафрази и на релации между медицински понятия (основно върху UMLS).

В Глава 3 се разглеждат средствата и методите за извличане на структурирани описания от медицински (български и английски) текстове. Изследвани са възможностите за разпознаване на събития в медицински текстове и на времевите аспекти в тях. Използвани са подходи, базирани на правила, и подходи от машинно самообучение.

В Глава 4 „Интеграция и прототипи” (неправилно посочена в автореферата като Глава 3) е демонстрирана интеграция на прототипи главно в областта на диабета и са извършени съответните експерименти за извличане на правила от положителни и отрицателни примери.

Приемам напълно изброените в дисертационния труд научни и научно-приложни приноси, представени от дисертанта. Представеният автореферат правилно отразява съдържанието и приносите на дисертационния труд.

Трябва да се отбележи точността и прецизността на автора при представяне на моделите и прототипите в дисертацията, тяхното формализирано описание, както и качеството на самия дисертационен труд.

Приложеният списък от 10 публикации на английски език по дисертацията подкрепят сериозната и задълбочена работа на дисертанта и показват неговото основно участие в съавторство. Публикациите удовлетворяват препоръчителните изисквания за образователната и научна степен "доктор". Даден е списък на 4 цитирания на публикации на дисертанта. Резултатите от дисертацията са докладвани на редица международни конференции, като докладът на AIMSА 2014 е получил награда за най-добра статия. Представено е и участието на дисертанта в редица национални и международни проекти, пряко свързани с дисертацията.

Нямам съществени критични бележки.

Представеният дисертационен труд съдържа значими резултати и решения на сериозни изследователски и приложни проблеми при обработката на естествен език и изграждането на семантични системи с приложения за клинични биомедицински текстове.

Авторът демонстрира отлично владение на професионалния инструментариум на теорията и практиката на редица информационни технологии при разработените от него методи, модели и приложни решения.

## ЗАКЛЮЧЕНИЕ

Всичко казано по-горе ми дава основание убедено да препоръчам на почитаемото Научно жури да присъди образователната и научна степен "Доктор" по научна специалност 01.01.12 „Информатика“ в професионално

направление 4.6 „Информатика и компютърни науки“ на Ивелина Мирчева  
Николова за дисертационния ѝ труд „Приложение на обработката на  
естествен език за изграждане на семантични системи”.

14.01.2015

София