

СТАНОВИЩЕ

на проф. д-р Мария Нишева – ФМИ, СУ „Св. Климент Охридски“
за дисертационен труд за придобиване на образователната и научна степен
“доктор” в професионално направление 4.6 „Информатика и компютърни
науки“
на тема „Лингвистични и семантични ресурси при компютърно генериране и
анотиране на български текстове“
с автор Каменка Атанасова Стайкова

Със заповед № 70/22.05.2015 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 специалност 01.01.12 „Информатика“, професионално направление 4.6 „Информатика и компютърни науки“, от Каменка Атанасова Стайкова с дисертационен труд на тема „Лингвистични и семантични ресурси при компютърно генериране и аотиране на български текстове“.

1. Обща информация за докторанта

Каменка Стайкова има висше образование по специалност „Информатика“, завършено през 1993 г. във Факултета по математика и информатика на СУ „Св. Климент Охридски“. Работила е като програмист, след което от 2000 г. заема академична длъжност в ИИТ, понастоящем – в ИИКТ при БАН. Участвала е в работата по редица национални и международни научноизследователски проекти.

Научните интереси и досегашните изследователски резултати на Каменка Стайкова са в областите обработка на знания, компютърна лингвистика, семантични технологии.

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

Дисертационният труд е посветен на изследвания в областта на компютърното генериране на текст. Негова цел е осъществяването на автоматично генериране на текстове на български език чрез създаване на лингвистични ресурси за основни обекти на българския език и на семантични ресурси за конкретни предметни области.

Дисертацията се състои от 180 страници текст, включващ увод, четири глави, заключение, списък на 83 използвани литературни източници и пет приложения.

Първа глава има обзорен характер. В нея е представена областта на изследванията – компютърното генериране на текст, както и са описани различните типове ресурси и технологии, които се използват в процеса на генериране.

Втора глава съдържа анализ на проблематиката, свързана със създаване на ресурси за генериране на изречения на български език. Описани са принципите на създаване на приложна българска системично-функционална граматика. Представени са резултати от автоматичното генериране на изречения на български език със създадената приложна българска системично-функционална граматика.

В трета глава се дискутира използването на приложната граматика в качеството на лексико-граматически ресурс за генериране на текстове на български език в конкретна предметна област. Описан е процесът на планиране на текст. Представени са резултати от автоматично генериране на текстове на български език в различни стилове.

Четвърта глава е посветена на ролята на семантичните технологии като ресурс за представяне на знания и семантично аотиране в процеса на автоматично генериране на текст. В рамките на конкретна предметна област е извършено изследване на процеса на семантично аотиране с частични граматика на базата на регулярни изрази.

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

- Създадено е формално описание на базови обекти от българския език, което подпомага автоматичното генериране на техническа документация на български език;
- Разработеното формално описание е реализирано във вид на разширяем и преизползваем компютърен ресурс – приложна системично-функционална граматика на българския език;
- Разработената приложна системично-функционална граматика на българския език е обособена като модул на многоезиковата система AGILE, която позволява компютърно генериране на технически текстове в различни стилове;
- Предложена е схема за семантично аотиране на специализирани текстове на български език.

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 поставената цел.

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

Резултатите, получени в дисертационния труд, са отразени в 12 статии, публикувани както следва:

- три – в списанието *Cybernetics and Information Technologies*;
- една – в списанието *Problems of Engineering Cybernetics and Robotics*;
- една – в списание „Български език“;
- шест – в сборници с доклади на международни научни форуми;
- една – във вътрешно издание на ИИТ – БАН.

Три от публикациите са самостоятелни, а девет са в съавторство. Нямам никакво съмнение в същественния личен принос на докторанта в колективните публикации.

Не са посочени цитирания на публикации на Каменка Стайкова по резултати от дисертационния ѝ труд, но при търсене с помощта на Google Scholar намерих седем цитирания на две от нейните статии, свързани с темата на дисертацията, както и множество цитирания на други нейни публикации.

4. Критични бележки и препоръки

Като недостатък на дисертационния труд бих посочила липсата на явно отделяне в глава 2 и глава 3 на собствените резултати на докторанта от евентуалните такива на други участници в проекта AGILE.

В текста на дисертацията са допуснати относително малък брой граматически грешки, които не оказват съществено негативно влияние върху общото впечатление от нейното съдържание.

5. Автореферат

Авторефератът отразява пълно и точно съдържанието и приносите на дисертационния труд, но обемът му е твърде голям. Списъкът на литературните източници съдържа такива, които не са цитирани в текста.

6. Заключение

Въз основа на изложеното по-горе смятам, че дисертационният труд напълно удовлетворява изискванията на ЗРАСРБ, ПЗРАСРБ и Правилника за специфичните

условия за придобиване на научни степени и за заемане на научни длъжности в ИИКТ при БАН. Убедено предлагам на уважаемото научно жури да присъди на автора му Каменка Атанасова Стайкова образователната и научна степен „доктор” в професионално направление 4.6 „Информатика и компютърни науки“.

София, 15.06.2015 г.