

# РЕЦЕНЗИЯ

за конкурса за професор по научната специалност 01.01.12 “Информатика (Езикови и семантични технологии)”, обявен в ДВ бр. 55 от 19.07.2011 г. за нуждите на ИИКТ-БАН, с единствен участник доц дмн Галя Младенова Ангелова

Рецензент: проф. д.м.н. д.т.н. Красимир Тодоров Атанасов

От представената професионална биография на кандидата се вижда, че тя е родена на 1.10.1955 г. През 1978 г. получава степен Магистър по математика, специалност Математическо осигуряване (шифър 01.01.12) във Факултета по математика и механика (сега – Факултет по математика и информатика) при Софийския университет “Климент Охридски”. През 1988 г. защитава като задочен докторант дисертация за придобиване на образователната и научна степен “Доктор” в Института по автоматизация и изчислителна техника на Унгарската академия на науките, под ръководството на акад. Янош Деметрович (трябва да отбележим, че дисертацията ѝ е издадена на руски език в поредицата *MTA SZTAKI Tanulmányok*, Будапеща). От 1980 г. до 1992 г. придобива последователно трите степени на “научен сътрудник” в ИМИ – БАН и КЦИИТ – БАН, от 1998 г. е ст.н.с. II ст. в ИПОИ – БАН, а от 2010 г. е „доктор на математическите науки”. От края на 2006 г. е ръководител на секция “Лингвистично моделиране” в ИПОИ. На практика през първите години на научната си кариера Г. Ангелова е работила в областта на базите от данни (в нея е получила образователната и научна степен “доктор”), а по-късно насочва интересите си към изкуствения интелект.

Доц. Г. Ангелова е автор общо на 116 научни труда. За участие в конкурса тя е представила 15 научни труда, 14 от които са статии в чуждестранни рецензирани издания, публикувани след 2003 г. Статии [1,4,5,7,8,12,13,14] са в серия “Lecture Notes in Computer Science / Artificial Intelligence” на Шпрингер, която до 2005 г. имаше „impact-factor”, нарастващ в последните години. За 2003 г., когато е публикувана статия [1], той е бил 0,251. Серията престана да има импакт-фактор по причини от извън-научен характер и затова приемам, че всичките 8 статии имат наукометричната тежест на статия в списание с импакт-фактор. Статии [11,15] са в серия “Studies in Health Technology and Informatics” на IOS Press с H-index 21. От всичките 15 статии, 14 са публикувани в списания или томовете от научни поредици на реномирани издателства, а една статия - [6] в сборник трудове на международна конференция. Всички те са пряко свързани с шифъра на конкурса – 01.01.12. Както е посочено в „Резюмето на научните постижения в статиите, представени при кандидатстване на конкурса за професор”, никой от представени трудове не е използван от доц. Ангелова в предишни процедури за научни степени или звания.

Приемам авторската справка за приносите в трудовете, които според кандидата се групират, както следва:

- Взаимодействие между езикови и семантични технологии в семантичния интернет или при построяване на декларативни семантични ресурси и дигитализирани архиви от учебни обекти (3 статии);

- Изучаване на поведението или построяване на приложения чрез статистически методи за обработка на естествен език (2 статии) , и
- Автоматична обработка на български текст в клинични записи на пациенти (10 статии).

По същество, резултатите на доц. Ангелова са научни и научно-приложни.

Публикациите от първата група са [1, 4, 9]. Първата от тях има полемичен характер и обяснява защо основният носител на информация в интернет е текстът на естествен език и не може да се построи интернет над знания. Илюстрират се предимствата на семантичния Уеб. В [4] се дискутират възможностите за приложение на модела LCIM (Levels of Conceptual Interoperability Model) към проблемите на семантичната оперативна съвместимост на учебни ресурси в семантичния Уеб. Статия [9] анализира състоянието на езиковите технологии, които могат да извлекат автоматично “концептуални” единици от технически текст. Изследва се коректността на Интернет-приложенията за анализ на естествен език.

Публикациите от втората група са [2, 3]. Статия [3] описва системата MorphoClass за автоматична обработка на немски текст, докато статия [2] е посветена на сравнителен анализ на различни техники за измерване на близост между документи. По мое мнение, тя е много интересна и може да стане база за бъдещо изследване, в което апаратът на интуиционистки размитите множества да се използва за оценяване на близостта между документите.

Основната част от публикациите на доц. Ангелова (10 статии) е свързана с автоматичната обработка на български текст в клинични записи за пациенти – статии [5 – 8, 10 – 15] представят резултати в областта на автоматичната текстообработка на медицински текстове. В статия [5] се изследват проблемите, които възникват при автоматичен анализ на медицински текстове, особено тези, които съдържат латински термини и специфични съкращения. За тази статия в авторската справка за приносите е посочено *„Представен е един основен принцип на дизайна на системата: за всеки пациент да се създаде един концептуален граф, който да бъде “налаган” към графите на други пациенти с цел търсене на подобие.”* Бих искал да предложа в бъдеще да се проучи възможността за тази цел да се използват и интуиционистки размити графи. Статии [6 – 8] се явяват естествено продължение на [5]. Те са богато илюстрирани с конкретни примери. В [10 – 15] са предложени алгоритми за нормиране на стойностите на признаците, по които се определя състоянието на пациентите и тяхното формализиране във вид на семантични мрежи, дискутира се идеята за автоматично определяне на отделните епизоди от заболяването на даден пациент на базата на описанието на развитието на болестта му, както и на нейното медикаментозно лечение. Даден е пример от областта на ендокринологията.

Както е посочено в авторската справка, изследванията са проведени над данни *„от клинични тестове от 6200 болнични записи на български език.”*. Като човек, имащ почти 30-годишен опит от работа с лекари, мога да потвърдя, че такива извадки от данни са изключително рядко срещани в дейността на български лекари. Този факт говори не само за упоритостта на екипа, който ги е обработил, но и за значим организационен опит на ръководителя доц. Г. Ангелова. Неслучайно, както е посочено и в авторската справка, *„тази статия получи Rolf Hansen prize за най-добра статия на*

Европейския конгрес по медицинска информатика в Осло през август 2011 г.” Пак в справката е посочено, че „статииите, публикувани през последните 1-2 години, показват и насоката за непосредствената бъдеща работа на Г. Ангелова, а именно:

- прецизиране на подхода за автоматично извличане на различни единици от текста на болнични записи,
- прилагане на създадените прототипни екстрактори към големи ресурси от болнични записи и
- създаване на методи, които се базират на автоматично извлечените показатели и осигуряват по-нататъшната им обработка с цел получаване на полезни медицински системи: за търсене на подобие между историите на заболяванията на различни пациенти, за намиране на странични ефекти на лекарства и т.н.

Мисля, че тази афиширана от доц. Ангелова програма за бъдещата ѝ дейност е много амбициозна и заслужава адмирация.

Доц. Г. Ангелова е предоставила списък от 147 цитирания, от които 52 цитирания са след защитата на дисертационния ѝ труд за „Доктор на математическите науки”. От справката, която направих със *scolar.google.com* установих, че тя има още доста нови цитирания, които не са отразени в документите ѝ. Например, за статия № 39 от списъка на публикациите, непредставени в конкурса, е посочено, че има 17 цитирания, докато аз открих 4 нови.

От приложената справка за учебната дейност на доц. Ангелова се вижда, че тя е била асистент по дискретна математика от 1979 до 1985 г., по компютърна лингвистика през 1987-89 г., а от 1997 г. чете изборни курсове „Увод в обработката на естествен език” и „Увод в компютърната лингвистика” във Факултета по математика и информатика при СУ.

Доц. Г. Ангелова е била научен ръководител на трима докторанта, от които двама вече успешно защитили и на 14 дипломанти, някои от които понастоящем вече са утвърдени учени, работещи в чужбина.

От приложените данни се вижда, че доц. Г. Ангелова е член на редакционната колегия на изданията на Международната конференция по концептуални структури ICCS (International Conference on Conceptual Structures), които се публикуват ежегодно в международното издателство Springer в серия Lecture Notes in Artificial Intelligence – от 2001 г. Тя е:

- член на НЕК по Информатика на Национален фонд „Научни изследвания”, 2001-2005 г.;
- национално контактното лице по ИКТ в 6-тата Рамкова програма и 7-мата Рамкова програма, с номинация от МОН/МОМН, от 2003 г. насам;
- член на ACL (Association of Computational Linguistics) и на Борда за Европа на ACL (EACL - 2003-2006 г.);
- член на Научния борд на Европейската инфраструктура КЛАРИН 2009-2010 г.;
- член на Технически комитет 14 на IFIP (2006 и 2007).

От 1992 г. насам доц. Ангелова е ръководител на научни проекти, тясно свързани с тематиката и на настоящия конкурс. Те са финансирани от Европейската комисия в

рамките на 5-та, 6-та и 7-ма рамкови програми, от ФНИ и от чуждестранни (немски и португалски) фондации.

Всичко това показва, че доц. Ангелова се проявява като добър организатор и администратор.

Според последния Закон за развитие на академичния състав в Република България, Правилника да прилагането му и Вътрешния правилник на ИИКТ – БАН, кандидатите за заемане на академичната длъжност **„Професор”** трябва да отговарят на следните:

1. Да са придобили образователната и научна степен „Доктор”;
2. Да са заемали академичната длъжност „доцент” не по-малко от пет години;
3. Да са представили публикуван монографичен труд или равностойни публикации в специализирани научни издания, които да не повтарят представените за придобиване на образователната и научна степен „доктор” и за заемане на академичната длъжност „доцент”;
4. Да са представили други оригинални научноизследователски трудове, публикации, изобретения и други научни и научно-приложни разработки, които се оценяват по съвкупност.

Доц. Г. Ангелова е не само „Доктор”, но и „Доктор на математическите науки”. Тя е „Доцент” от 1998 г. За конкурса тя е представила 15 статии, 10 от които са тясно свързани по тематика. По мое мнение, техният обем и съдържание са достатъчни, за да бъдат събрани в монография. Затова препоръчвам на доц. Ангелова в близко бъдеще наистина да я подготви. Следователно кандидатката напълно удовлетворява изискванията на цитираните по-горе документи за придобиване на академичната длъжност „Професор”. Освен това, нейните публикации и цитирания удовлетворяват и изискванията на бившите научни комисии на ВАК – за математически науки и за електротехнически науки, електроника и автоматика, които приемаха процедурите с шифър 01.01.12 „Информатика”, след като те са били разисквани и гласувани съответно в СНС по информатика и математическо моделиране и в СНС по електронна и компютърна техника (единствени СНС, имащи компетенции по този шифър).

Точно преди две години бях рецензент на дисертационния труд на доц. Ангелова за „Доктор на математическите науки”. Тогава написах, че *„познавам Г. Ангелова още от студентските ни години. За нея още от тогава се говореше, че е многообещаващ специалист, твърдение, която тя успешно защитава и до днес.”*

Казаното по-горе е основание да дам положителна оценка на материалите, с които кандидатът участва в конкурса за „професор” и да препоръчам на уважаемите членове на Научното жури, назначено със Заповед № 94/26.09.2011 г. на Директора на ИИКТ-БАН, а впоследствие и на уважаемите членове на Научния съвет на ИИКТ-БАН, да гласуват за присъждането на доц дмн **Галя Младенова Ангелова** на академичната длъжност **“Професор”** по научната специалност 01.01.12 **“Информатика (Езикови и семантични технологии)”**.

2. XI. 2011 г.