

РЕЦЕНЗИЯ

на доц. д-мн Галя Младенова Ангелова
във връзка с
конкурса за професор по 01.01.12. Информатика (Езикови и семантични
технологии), професионално направление 4.6 «Информатика и компютърни
технологии», обявен в Държавен вестник, бр. 55/19.07.2011 г.

от

доц. д-р Кирил Иванов Симов,
Секция Лингвистично моделиране,
Институт по информационни и комуникационни технологии,
Българска академия на науките

Галя Ангелова е родена на 01.10.1955 г. Тя завършва Факултета по математика и механика на СУ "Св. Климент Охридски" през 1978 г. Става доктор по математика през 1988 г. и доктор на математическите науки през 2010 г. Галя Ангелова е избрана за доцент по информатика в Централната лаборатория по паралелна обработка на информацията през 1996 г. Преподавателската ѝ дейност в периода 1979-2010 г. е била във СУ „Св. св. Кирил и Методий“, където Галя Ангелова е била асистент по дискретна математика и лектор по база данни, компютърна лингвистика и обработка на естествените езици. Имала е 15 дипломанти и докторанти. Има двама защитили докторанти. Участвала е в организирането на редица международни конференции и семинари, като най-значимата е конференцията RANLP (Recent Advances in NLP). Трябва да се отбележи, че RANLP получи статута на единствената българска конференция в областта на обработките на естествен език, която е изключително значима и която се утвърди като част от добрите практики в международен план. Основната заслуга за това е на Галя Ангелова (съвместно с професионалната ѝ екипна работа с Руслан Митков, Калина Бончева и колегите от ИИКТ).

Галя Ангелова е участвала и ръководила 14 български и международни проекта. Тя е рецензент за Европейската комисия, била е национално контактното лице в програмите по информационни технологии IST FP 6 и IST FP7, член на борда на Европейската асоциация по компютърна лингвистика, член на Експертния съвет по информатика на НФНИ-България, член на Специализирания научен съвет по информатика и математическо моделиране на ВАК, член на Научния борд на инфраструктурата КЛАРИН. Всички посочени дейности говорят за голямата работоспособност и чувство за научна отговорност на кандидатката в областта на нейната експертиза. Доказателство за умението да се комбинират научните способности с мениджърските умения е фактът, че Галя Ангелова е избрана единодушно за ръководител на Секцията по лингвистично моделиране от 2006 г. насам.

Галя Ангелова има над 115 публикации в научни списания и сборници, от които 15 са представени за конкурса. Също така има над 150 известни на авторката цитирания. Галя Ангелова е била поканен лектор в редица университети в чужбина. Не на последно място трябва да се споменем и научните контакти на кандидатката с научни институции и бизнес компании в България и чужбина, които показват, че тя е търсен специалист в областите, в които е работила до момента.

Галя Ангелова има редица важни приноси в областта на компютърната лингвистика и обработката на естествен език и в областта на семантичните технологии. Научната ѝ работа представя последователен анализ и имплементация на модели на диалога, генерирането на естествен език, моделирането на концептуалното знание и анализ на медицински текстове. Галя Ангелова се интересува както от приложни аспекти на езиковите технологии, така и от теоретичното моделиране на понятия, значението и начините на изграждане на термини, взаимодействието между езиковите и семантичните технологии при извличането на информация от текста и при изграждането на семантичния уеб.

В областта на семантичния уеб Галя Ангелова има приноси при семантичното унифициране на текстови ресурси с оглед на тяхното съвместно използване чрез съпоставка към обща онтология (в представената публикация в областта на учебните обекти). Друг принос на Галя Ангелова в тази област е на извличане на концептуално знание от текстове на базата на термини, използвани технически термини. Тук Галя Ангелова е прилагала различни подходи за по-пълно разпознаване и групиране на намерените термини като термини от повече от една дума, онтологични връзки между термини в текста, използване на лингвистично знание за клъстеризация на термини по семантични признаци. Галя Ангелова също така има принос към създаването на онтологии, семантичното аотиране на текстове с такива онтологии и използването им при семантично търсене.

В областта на прилагане на статистически методи при обработката на естествен език Галя Ангелова се интересува от намирането на близост между документи и класификацията на непознати думи. Основният използван метод е Latent Semantic Analysis (LSA), който разглежда документите като вектори от езикови единици и чрез Singular Value Decomposition се оптимизира големината на разглежданото пространство. Галя Ангелова има принос при определяне на оптималните обработки на текстовете на документите с цел да се постигне най-добро определяне на близостта на документи. Използвани са анализи на ниво корен, основна форма на думите, частични граматика, използване на стоп-думи и други. Методът LSA е сравнен с други методи за определяне на близост на документи. Методът разработен за системата MorphoClass показва много добри резултати и е възможно да се адаптира за други флективни езици.

В областта на автоматичната обработка на български текстове от клинични записи на пациенти Галя Ангелова е пионер сред изследователите в България. Разработена е система за автоматично извличане на информация за пациенти, тяхното състояние, назначеното лечение и лекарствата, с които са лекувани. Задачата е особено трудна поради спецификата на текста, който изобилства със съкращения, латински термини, елипси и фрагменти. Този тип текстове изисква

създаването на много модули за обработка на езика от нула, т.к. модулите за стандартния език са неприложими. Тази област на приложение на езиковите и семантичните технологии е предизвикателство поради голямото количество концептуално знание, многобройните законови разпоредби и естеството на приложение. За да може ефективно да се извличат необходимите факти от текстовете е необходимо голямо по обем знание за областта, което да поддържа системата при запълването на липсващите понятия и факти. Галя Ангелова успешно е интегрирала в своите разработки методи като онтологии за представяне на концептуалното знание, концептуални графи за представяне на извлечените факти, евристични методи за откриването на фактите в документа. За качеството на разработките в тази посока говорят както отзивите на потребители им в България, така и оценката на статията „Boycheva, S., D. Tcharaktchiev and G. Angelova. Contextualization in automatic extraction of drugs from Hospital Patient Records. In Proc. of MIE-2011, Oslo, Norway.” като най-добра статия на конференцията.

Не по-малко важни са приносите на кандидатката и в областта на обучението. От една страна, тя е дългогодишен преподавател по Компютърна лингвистика във ФМИ, където е успяла да ориентира редица студенти за развитие в тази област. Днес някои от тях изявиени изследователи, като например Преслав Након, Кристина Тутанова, Ани Ненкова, или успешни директори на фирми за информационни технологии, като Атанас Киряков и др. От друга страна, Галя Ангелова е участвала активно в създаването на учебни материали за чуждоезиково обучение чрез прилагане на съвременните методи за обработки на езика. Също така, чрез европейския проект BALRIC-Ling, на който Галя Ангелова беше координатор, беше създадена информационна база с материали в областта на компютърната и корпусната лингвистика – за корпуси, морфологични и синтактични анализатори, за различните видове електронни речници и т.н. Тези материали се използват активно в курсовете по компютърна лингвистика в СУ „Св. Кл. Охридски”.

Галя Ангелова има важни приноси в областта на езиковите технологии и представянето на знания. Тя прилага иновативни подходи при анализ на текст и извличане на информация с оглед на различните задачи. Галя Ангелова е пример за учен, който балансира теоретичните изследвания с практическата приложимост в реални контексти. Тя има визия за развитието на научната си област не само по отношение на краткосрочните цели, но и в дългосрочна перспектива.

Въз основа на посочените научни приноси и научната автобиография на Галя Ангелова, давам висока оценка на постиженията ѝ и убедено препоръчвам на уважаемото Научно жури да я избере за професор по шифър 01.01.12. Информатика (Езикови и семантични технологии). Смятам, че Галя Ангелова отговаря на всички необходими критерии, за да получи това звание.

30.11.2011