

СТАНОВИЩЕ

по процедура за защита на дисертационен труд на тема:

Автоматично резюмиране на правни текстове

за придобиване на образователна и научна степен „доктор“ от

кандидат: **Валентин Венков Змийчаров,**

Област на висше образование: **4. Природни науки, математика и информатика**

Професионално направление: **4.6. Информатика и компютърни науки**

Докторска програма: **„Софтуерни технологии - Откриване на знания“,**

катедра: **„Софтуерни технологии“,**

Факултет по математика и информатика (ФМИ),

Софийски университет „Св. Климент Охридски“ (СУ)

Становището е изготвено от **проф. д-р Олга Илиева Георгиева – ФМИ СУ** в качеството ми на член на научното жури, съгласно Заповед № РД 38-278/ 9.06.2025 г. на Ректора на Софийския университет.

1. Обща характеристика на дисертационния труд и представените материали

Представеният за становище дисертационен труд се състои от 114 страници, обхващащи увод, четири глави по същество и една заключителна глава, литературна справка от 58 източника, списък от 27 фигури и списък от 9 таблици. Представеният обем, съдържание, структура на докторската теза покриват критериите за такъв научен труд.

Материалите, предоставени от кандидата, за оформяне на становището съдържат документи, които дават пълна информация за кандидатурата. Освен дисертационния труд, са предоставени още: резюме на български и английски език, публикациите по труда, също дипломи, удостоверения, справки, доклади, становища в подкрепа на докторантурата. Всички те уверяват, че изискванията на ЗРАСРБ и Правилника за неговото приложение, както и Правилника на СУ за условията и реда за придобиване на научни степени и заемане на академични длъжности, са спазени.

2. Данни и лични впечатления за кандидата

Валентин Змийчаров е завършил бакалавърска степен със специалност „Софтуерно инженерство“ и магистърска със специалност „Извличане на информация и откриване на знания“ и двете придобити във ФМИ на СУ. Като негов преподавател имам преки и отлични впечатления от представянето му и в двете степени на обучение – активен, точен и трудолюбив студент. Допълнителна пряка представа от работата му придобих от времето на неговата задочна докторантура към катедра „Софтуерни технологии“. В рамките на годишните атестации съм имала възможност да проследя развитието на работата по представения дисертационен труд.

Допълнително, от автобиографичната справка на кандидата прави впечатление широкия обхват на придобити знания и умения, както и добрата му професионална реализация в областта на софтуерното инженерство.

3. Съдържателен анализ на научните и научно-приложните постижения на кандидата, съдържащи се в представения дисертационен труд и публикациите към него, включени по процедурата

Резюмирането на текстове и в частност тези, съдържащи правна информация, което е научна цел на дисертационния труд, е от особено значение в съвременния живот на обществото. Автоматичното генериране на резюмета е начин за подобряване на ефективността на работата в сектора, също възможност за постигне нужната безпристрастност и коректност в използваната информация. В изпълнение на поставената цел дисертационният труд демонстрира научни постижения, които обобщено могат да бъдат представени както следва:

Научни резултати: Разработен е нов модел - StructExtSum, използващ структурирана информация на документа, за ефективно извличане на релевантни пасажии и оформяне на резюме на документа.

Научно-приложни резултати: За целите на резюмиране на правни текстове са адаптирани известни екстрактивни подходи – Term Frequency–Inverse Document Frequency (TF-IDF), TextRank и подходи, основани съответно на BERT и LegalBERT модели, както и три абстрактивни модела за резюмиране –PreSumm, езиков модел T5, езиков модел GPT. Представен е сравнителен анализ на резултатите от приложените подходи чрез аналитични оценки за сходство, както и на база на мнение на оценители на формулираните резюмета.

Приложни резултати: Създаден е набор данни, съдържащ юридически текстове на 24 езика и съответни резюмета, оформени от експерти в областта. Създадената база данни е публикувана в платформата Hugging Face.

4. Аprobация на резултатите

В подкрепа на научните резултати от дисертационния труд кандидатът е представил 4 публикации в сборници на международни научни конференции в областта на компютърните науки. Две от конференциите са тясно специализирани в тематиката на компютърната обработка на естествен език.

Считам, че публикациите са добре представени на научната общност. Четирите публикации са видими в системата SCOPUS и имат SJR индекс, а една статия (тази от 2024 година) - и в системата WoS. Така те събират общо 120 точки според минималните национални изисквания на ЗРАСРБ, което е значително над изискуемия минимум. Публикациите са в съавторство като в три от тях кандидатът е на първо място. За всички тях са представени подписани декларации за равностойно участие на авторите. Към деня на изготвяне на рецензията три цитирания (без автоцитати) на публикации по дисертационния труд са видими в SCOPUS.

Направената справка позволява категорично да се обобщи следното:

а) Научните трудове отговарят на минималните национални изисквания (по чл. 2б, ал. 2 и 3 на ЗРАСРБ) и съответно на допълнителните изисквания на СУ „Св. Климент Охридски“ за придобиване на образователна и научна степен „доктор“ в научната област и професионално направление на процедурата;

б) Представените от кандидата резултати в дисертационния труд и научни трудове към него не повтарят такива от предишни процедури за придобиване на научно звание и академична длъжност. За кандидата настоящата процедура е първа за придобиване на научно звание и той не е участвал в процедури за заемане на академична длъжност.

в) Няма доказано по законоустановения ред плагиатство в представения дисертационен труд и научни трудове по тази процедура.

5. Качества на автореферата

Авторефератът на дисертационния труд е оформен в 45 страници на български език и 41 страници на английски език. Неговото съдържание следва представените изследвания в дисертационния труд като точно са описани приносните части в

отделните глави. Авторефератът обобщава пълно, коректно и в достатъчна степен представя анализите и резултатите на научния труд.

6. Критични бележки и препоръки

Получените научни резултати са оригинални и след допълване и разширяване биха могли да бъдат представени за публикуване както на научни конференции в чужбина, така и в международни научни списания.

7. Заключение

След като се запознах с представените в процедурата дисертационен труд и придружаващите го научни трудове и въз основа на направения анализ на тяхната значимост и съдържащи се в тях научни, научно-приложни и приложни приноси, **потвърждавам**, че представеният дисертационен труд и научните публикации към него, както и качеството и оригиналността на представените в тях резултати и постижения, отговарят на изискванията на ЗРАСРБ, Правилника за приложението му и съответния Правилник на Софийски университет „Св. Климент Охридски“ за придобиване от кандидата на образователната и научна степен „доктор“ в научната област 4. Природни науки, математика и информатика и професионално направление 4.6. Информатика и компютърни науки. В частност кандидатът удовлетворява минималните национални изисквания в професионалното направление и не е установено плагиатство в представените по конкурса научни трудове.

Въз основа на гореизложеното, **препоръчвам** на научното жури да присъди на **Валентин Венков Змийчаров** образователна и научна степен „доктор“ в научна област 4. Природни науки, математика и информатика, професионално направление 4.6. Информатика и компютърни науки.

6.08.2025 г.

Изготвил становището:

(проф. д-р Олга Георгиева)