

## ВЕЋУ ЗА СТУДИЈЕ ПРИ УНИВЕРЗИТЕТУ У БЕОГРАДУ

На редовној седници Већа за студије при Универзитету у Београду одржаној 23. марта 2026. године, **Одлуком 06 Број: 06-4574/IV/971/2-26 JKJ/** именован/а сам за члана Комисије за оцену научне заснованости предложене теме докторске дисертације под насловом **Мултиагентски приступи засновани на великим језичким моделима за објашњиво аутоматско оцењивање студентских радова** и испуњености услова кандидата **Марка Путниковића** и предложених ментора **др Јелене Јовановић** и **др Александра Бауцала**.

На основу поднете документације уз Пријаву теме докторске дисертације кандидата Марка Путниковића, подносим Већу за студије при Универзитету у Београду следећи:

### ИЗВЕШТАЈ О ОЦЕНИ НАУЧНЕ ЗАСНОВАНОСТИ ТЕМЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ И ИСПУЊЕНОСТИ УСЛОВА КАНДИДАТА И МЕНТОРА

#### 1. Биографија кандидата

Марко Путниковић рођен је 13.5.1993. године у Ваљево, Србија. Уписао је основне студије на Факултету организационих наука Универзитета у Београду 2013. године, смер Информациони системи и технологије. Дипломирао је 2017. године стекавши звање Инжењер организационих наука - информациони системи и технологије. Мастер студије је уписао 2017. године на Универзитету у Греноблу, на смеру Вештачка интелигенција и семантички веб, као стипендиста владе Француске. Мастер студије је завршио 2018. године стекавши звање Мастер информатичар. Током мастер студија је обавио петомесечну праксу (stage) при Лабораторији за информатику Универзитета у Греноблу. Докторске студије је уписао 2019. године на Универзитету у Београду на програму мултидисциплираних студија при универзитету, смер Интелигентни системи. Током студија се бавио применом метода процесирања природних језика у области едукације и објавио истраживачки рад у M21 часопису о примени тих метода у оцењивању студентских радова (2.1.1, 2.2.1).

Упоредо са докторским студијама, кандидат је радио у компанији NCR као софтвер инжењер 2019-2022. године, након чега прелази у компанију Robert Bosch у којој и даље ради као инжењер вештачке интелигенције. У компанији Robert Bosch је радио на иновационом истраживачком пројекту о примени метода процесирања природних језика за екстракцију података са веба, у оквиру ког је објавио истраживачки рад на конференцији (2.2.2).

#### 2. Библиографија кандидата (категорисано према категоризацији надлежног Министарства)

##### 2.1. Објављени радови или прихваћени за штампу (DOI број или изјава уредника)

Истакнути међународни часописи (M21):

- 2.1.1. Putnikovic, M., & Jovanovic, J. (2023). Embeddings for automatic short answer grading: A scoping review. *IEEE Transactions on Learning Technologies*, 16(2), 219-231. <https://doi.org/10.1109/ilt.2023.3253071>

##### 2.2. Конференције и скупови

- 2.2.1. Putnikovic, M., & Jovanovic, J. (2022). Word embeddings for automatic short answer grading. In *Digital Transformation for the University of the Future* (pp. 197-213). World Scientific Publishing Co. Pte. Ltd. [https://doi.org/10.1142/9789811254154\\_0011](https://doi.org/10.1142/9789811254154_0011)
- 2.2.2. Häty, A., Milchevski, D., Döring, K., Putnikovic, M., Mesgar, M., Novović, F., ... & Stranjanac, I. (2024). A cost-efficient modular sieve for extracting product information from company websites. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track* (pp. 1444-1456). <https://doi.org/10.18653/v1/2024.emnlp-industry.106>

- 2.2.3. Stanković, R., Rađenović, J., Škorić, M., & Putnikovic, M. (2025). Learning word embeddings using lexical resources and corpora. *ICIST - International Conference on Information Society and Technology (ICIST 2025)*, Kopaonik, Serbia. <https://doi.org/10.5281/zenodo.15093900>

### 3. Предмет и циљеви докторске дисертације

#### 3.1 Предмет докторске дисертације (максимално 1 страна)

Предмет докторске дисертације је аутоматско оцењивање студентских радова и пружање објашњења о процесу оцењивања применом мултиагентских система заснованих на великим језичким моделима.

Оцењивање студентских радова представља један од главних процеса у оквиру образовног система, јер омогућује процену и пружа повратну информацију о усвојености знања студента. Оцењивање може бити сумативне природе, што подразумева доделу оцене и даје формалну основу за напредовање и стицање академског звања; и формативне природе, које укључује пружање повратне информације студенту, праћење напретка у учењу и усмеравање даљег развоја компетенција студента. За наставника, задатак оцењивања је обично временски ограничен и когнитивно напоран у зависности од броја студената, што може довести до замора, утицати на доследност и смањити могућност боље подршке студентима. Додатно, субјективност наставника, тачније, лична уверења и очекивања, ослањање на претходно искуство и интуицију може довести до неконзистентности при оцењивању и утицати на исход. Са развојем онлајн облика образовања, све већи значај добијају системи за аутоматско оцењивање, чија је намена да се омогући брже и конзистентније оцењивање великог броја студентских радова, као и да наставнике растерети дела процеса оцењивања и пружања образложења додељене оцене студенту који се може аутоматизовати.

Формално, задаци који се оцењују се могу поделити на оне затвореног типа (нпр. задатак са више понуђених одговора) и отвореног типа (конструисан одговор) (Bennet, 1991). Задаци затвореног типа подразумевају одабир једног или више понуђених одговора, при чему је оцењивање детерминистичко, тј. засновано на унапред дефинисаном решењу. С друге стране, задаци отвореног типа се могу поделити на оне са кратким и дугим текстом слободног одговора (есеји), али и у зависности од домена предмета могу бити програмски код, математички доказ и сл. За отворен тип одговора, који ће бити фокус истраживања ове докторске дисертације, потребна је обрада и анализа садржаја, што захтева примену софистициранијих метода процесирања природних језика.

Историјски, методе оцењивања задатака отвореног типа се могу поделити на оне (1) засноване на шаблонима и правилима, (2) статистичким карактеристикама текста, и (3) машинском и дубоком учењу (Burrows et al. 2015). Приступ заснован на дубоком учењу је посебно присутан последњих година услед значајног успеха примене великих језичких модела на оцењивање студентских радова (Fagbohun et al. 2024). Међутим, у великом броју постојећих истраживања све задатке у оквиру процеса оцењивања обавља један велики језички модел, што може довести до разних проблема који се генерално приписују таквим моделима, као што су халуцинације и губитак контекста. Осим тога, оцењивање једним великим језичким моделом може довести до мешања критеријума за оцењивање уколико је потребно оценити различите аспекте рада понаособ. Додатно, информације које један велики језички модел поседује могу бити застареле или нерелевантне, те је потребно обогатити упит текућим знањем битним за контекст оцењивања.

Приступ којим би се могао унапредити процес оцењивања и пружања објашњења тог процеса помоћу великих језичких модела јесте примена система оцењивања заснованих на више језичких модела у међусобној спреси. Такви системи се називају мултиагентским системима (МАС), где је агент софтверска компонента која, између осталог, користи велики језички модел како би обавила одређени задатак у процесу оцењивања (Kamalov et al. 2025). Поред великих језичких модела, агент може користити разне друге програмске алате како би обогатио своје текуће знање о релевантном контексту и дао бољи резултат (претрага веба, библиотеке за анализу података, приступ базама података и сл). Примена МАС-а заснованих на великим језичким моделима у аутоматском оцењивању и даље представља релативно нову и недовољно систематски истражену област. Посебно недостају истраживања која анализирају утицај архитектуре, тј. структурне организације МАС-а на резултате оцењивања. Конкретно, ограничен је број истраживања која испитују предности и недостатке мултиагентских приступа у односу на друге системе аутоматског оцењивања, испитују утицај специјализације агената по различитим аспектима оцењивања на тачност оцењивања, начине агрегације оцена појединачних агената у коначну оцену, и анализирају ефекте интеграције спољног знања у МАС. Поред тога, постојећа истраживања се недовољно баве питањем транспарентности и објашњивости процеса оцењивања у МАС-у. Образложење оцене коју додељује аутоматски систем неопходно је како би се обезбедила транспарентност оцењивања и омогућила провера примене рубрика за оцењивање, чиме се гради поверење у систем. Додатно, регулаторни правни оквири попут Акта о вештачкој интелигенцији Европске уније (European Union, 2024, Article 86) захтевају одређени ниво транспарентности аутоматског одлучивања, посебно у доменима примене са високим ризиком у које се убраја и образовање.

### 3.2. Циљеви докторске дисертације (максимално 1 страна)

Општи циљ ове докторске дисертације је развој тачног, конзистентног и објашњивог МАС-а заснованог на великим језичким моделима за аутоматско оцењивање студентских радова. У складу са дефинисаним општим циљем, формулисани су следећи конкретни циљеви:

1. Испитати у којој мери специјализација агената у односу на рубрику доприноси тачности аутоматског оцењивања
2. Испитати допринос интеграције релевантног спољног знања (нпр. курикулума или наставних материјала) на тачност аутоматског оцењивања
3. Анализирати како различити начини агрегације процена појединачних агената утичу на тачност аутоматског оцењивања
4. Обезбедити транспарентност и објашњивост процеса оцењивања МАС-а кроз генерисање утемељених и аргументованих објашњења

### 3.3 Хипотезе

На основу анализе релевантне и расположиве литературе, као и учених недовољно истражених тема у области аутоматског оцењивања, формулишемо следеће хипотезе:

- X1. Примена МАС-а у аутоматском оцењивању студентских радова доводи до повећања тачности оцењивања у односу на постојеће приступе.
- X2. Адаптивна специјализација агената у односу на тип задатка, заснована на рубрици за оцењивање, доводи до повећања тачности аутоматског оцењивања у односу на мултиагентски приступ без адаптивне специјализације агената.
- X3. Интеграција релевантног спољног знања у МАС, одређена типом задатка који је предмет оцењивања, доводи до веће тачности аутоматског оцењивања, у односу на МАС без употребе спољног знања.
- X4. Увођење агента за агрегацију у МАС, који на основу процена појединачних агената доноси коначну оцену, доводи до веће тачности аутоматског оцењивања у односу на МАС у којем се коначна оцена добија једноставном агрегацијом (нпр. упросечавањем) оцена појединачних агената.
- X5. Мултиагентски систем заснован на великим језичким моделима омогућује већу транспарентност и објашњивост процеса аутоматског оцењивања у односу на моноагентски приступ, кроз генерисање образложења донете оцене.

### 4. План рада

План рада на докторској дисертацији ће се састојати из следећих корака:

- преглед релевантне литературе из области аутоматског оцењивања, великих језичких модела и МАС-а
- анализа постојећих примена МАС-а заснованих на великим језичким моделима за аутоматско оцењивање
- избор јавно доступних скупова података за евалуацију аутоматског оцењивања есеја и кратких одговора, као и припрема скупа података за оцењивање програмског кода
- Прикупљање рубрика за оцењивање по типу задатка прегледом релевантне литературе и анализом рубрика из јавно доступних скупова података
- дизајн моноагентског система за аутоматско оцењивање (базни модел)
- дизајн МАС-а са адаптивном специјализацијом агената
- дизајн МАС-а са агентом за агрегацију оцена
- дизајн МАС-а са приступом спољном знању
- укључивање механизма за генерисање утемељених и аргументованих образложења оцене у оквиру МАС-а ради обезбеђивања транспарентност и објашњивости процеса оцењивања
- одабир репрезентативних великих језичких модела који ће се користити у експерименталној евалуацији
- дефинисање метрика за евалуацију, тј. процену тачности дизајнираних МАС-а, као и процену квалитета аутоматски генерисаних образложења оцена
- имплементација предложених система и експерименталних поставки
- извођење експеримената и прикупљање резултата
- анализа резултата, визуелизација и интерпретација у односу на постављене хипотезе
- дискусија ограничења истраживања и предлог будућих праваца развоја

### 5. Библиографски подаци релевантни за докторску дисертацију (максимално 1 страна)

Bennett, R.E. (1991). On the meanings of constructed response. *ETS Research Report Series, 1991*.  
<https://doi.org/10.1002/j.2333-8504.1991.tb01429.x>

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Burrows, S., Gurevych, I. & Stein, B. The eras and trends of automatic short answer grading. *Int J Artif Intell Educ* 25, 60–117 (2015). <https://doi.org/10.1007/s40593-014-0026-8>
- Chu, Z., Wang, S., Xie, J., Zhu, T., Yan, Y., Ye, J., ... & Wen, Q. (2025). LLM agents for education: Advances and applications. In C. Christodoulopoulos, T. Chakraborty, C. Rose, & V. Peng (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2025* (pp. 13782–13810). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.findings-emnlp.743>
- European Union. (2024). Artificial intelligence act (Regulation (EU) 2024/1689).
- Fagbohun, O., Iduwe, N. P., Abdullahi, M., Ifaturoti, A., & Nwanna, O. M. (2024). Beyond traditional assessment: Exploring the impact of large language models on grading practices. *Journal of Artificial Intelligence, Machine Learning and Data Science*, 2(1), 1-8. <https://doi.org/10.51219/jaimld/oluwole-fagbohun/19>
- Gao, M., Li, Y., Liu, B., Yu, Y., Wang, P., Lin, C. Y., & Lai, F. (2025). Single-agent or multi-agent systems? Why not both?. *arXiv preprint arXiv:2505.18286*.
- Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N. V., Wiest, O., & Zhang, X. (2024). Large language model based multi-agents: A survey of progress and challenges. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)* (pp. 8048–8057). <https://doi.org/10.24963/ijcai.2024/890>
- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in education: Principles, policy & practice*, 4(3), 365-379. <https://doi.org/10.1080/0969594970040304>
- Idowu, J. A., & Almasoud, A. (2026). Specialists or generalists? Multi-agent and single-agent LLMs for essay grading. *arXiv preprint arXiv:2601.22386*.
- Jiang, Y. H., Li, R., Zhou, Y., Qi, C., Hu, H., Wei, Y., ... & Wu, Y. (2024). AI agent for education: Von Neumann multi-agent system framework. *arXiv preprint arXiv:2501.00083*.
- Jiang, Y. H., Liu, T. Y., Zhuang, X., Hu, H., Li, R., & Jia, R. (2024). Enhancing educational practices with multi-agent systems: A review. *Enhancing Educational Practices: Strategies for Assessing and Improving Learning Outcomes*, 47-65. <https://doi.org/10.52305/ruig5131>
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Kamalov, F., Calonge, D.S., Smail, L., Azizov, D., Thadani, D.R., Kwong, T., & Atif, A. (2025). Evolution of AI in Education: Agentic workflows. *ArXiv, abs/2504.20082*.
- Lagakis, P., & Demetriadis, S. (2024, June). EvaAI: A multi-agent framework leveraging large language models for enhanced automated grading. In *International Conference on Intelligent Tutoring Systems* (pp. 378-385). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-63028-6\\_32](https://doi.org/10.1007/978-3-031-63028-6_32)
- Li, G., Chen, L., Tang, C., Švábenský, V., Deguchi, D., Yamashita, T., & Shimada, A. (2025, June). Single-agent vs. multi-agent LLM strategies for automated student reflection assessment. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 300-311). Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-96-8186-0\\_24](https://doi.org/10.1007/978-981-96-8186-0_24)
- Nitko, A. J., & Brookhart, S. M. (2006). *Educational assessment of students*. Prentice-Hall.
- Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation*. Routledge. <https://doi.org/10.4324/9780203122761>
- Wang, Y., Ding, Z., Wu, X., Sun, S., Liu, N., & Zhai, X. (2026, March). Autoscore: Enhancing automated scoring with multi-agent large language models via structured component recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 40, No. 48, pp. 40898-40906). <https://doi.org/10.1609/aaai.v40i48.42123>
- Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Awadallah, A., White, R. W., Burger, D., & Wang, C. (2024). AutoGen: Enabling next-gen LLM applications via multi-agent conversation. In *COLM 2024*.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?. *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhang, R. (2025, July). Multi-agent systems for learning assessment in education: A comprehensive survey. In *Proceedings of the 2025 3rd International Conference on Educational Knowledge and Informatization* (pp. 388-391). <https://doi.org/10.1145/3765325.3765390>

Zhang, X., Zhang, C., Sun, J., Xiao, J., Yang, Y., & Luo, Y. (2025). EduPlanner: LLM-based multi-agent systems for customized and intelligent instructional design. *IEEE Transactions on Learning Technologies*. <https://doi.org/10.1109/tlt.2025.3561332>

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*, 1(2), 1-124.

## 6. Методе које се користе у истраживању (максимално 1 страна)

Методологија истраживања биће усмерена на остваривање општег циља дисертације који укључује дизајн, развој и експерименталну евалуацију МАС-а заснованог на великим језичким моделима, за тачно, конзистентно и објашњиво аутоматско оцењивање студентских радова. У складу са тим, експерименти ће бити дизајнирани тако да омогуће проверу постављених хипотеза, као и остваривање наведених конкретних циљева дисертације.

У оквиру циља испитивања утицаја мултиагентског приступа на тачност оцењивања, а који се односи на хипотезу Х1, биће имплементирани: базни, моноагентски систем, и МАС за аутоматско оцењивање, оба заснована на великим језичким моделима. Поређење ова два система вршиће се на јавно доступним скуповима података за оцењивање есеја и кратких одговора, као и на припремљеном скупу података за оцењивање програмског кода. Поред међусобног поређења, резултати предложених система биће упоређени и са постојећим системима за аутоматско оцењивање који су на истим јавним скуповима података показали високе резултате, како би се у анализу укључили и приступи који нису засновани на агентским архитектурама.

Ради испитивања утицаја адаптивне специјализације агената на тачност аутоматског оцењивања, у складу са хипотезом Х2, биће имплементиран МАС са специјализованим агентима. Агенти ће бити специјализовани према појединачним критеријумима оцењивања дефинисаним у рубрици или утврђених претходним прегледом релевантне литературе. Адаптивност МАС-а са специјализацијом биће омогућена увођењем оркестратор агента који ће, на основу анализе рубрике за оцењивање за дати тип задатка, утврдити по којим критеријумима је потребно оценити студентски рад и на основу њих иницијализовати одговарајући број специјализованих агената. Поређење ће се вршити са МАС-ом без адаптивне специјализације агената, над истим скуповима података, чиме ће се утврдити у којој мери специјализација агената у односу на тип задатка утиче на тачност аутоматског оцењивања.

Како би се испитао утицај начина интеграције релевантног спољног знања на тачност оцењивања, у оквиру хипотезе Х3, биће имплементирани различити приступи коришћења спољног знања у оквиру МАС-а. У ту сврху биће коришћени разни извори знања, која могу бити општи за дати задатак (нпр. курикулум, уџбеник, референтни одговор) или специфични за критеријум оцењивања (нпр. дефиниције појмова, доменско знање). Код општег начина коришћења спољног знања, сви агенти ће користити заједнични општи контекст прибављен из општих извора знања. Код специфичног начина, агенти ће прибављати и користити знање релевантно свом критеријуму оцењивања. Поред тога, биће испитан и хибридни приступ, у оквиру ког ће агенти користити комбинацију општег и специфичног знања. Поређењем перформанси ових система међусобно, као и са верзијом МАС-а без приступа спољном знању, биће анализирано у којој мери интеграција и различити начини интеграције релевантног спољног знања утичу на тачност аутоматског оцењивања.

У циљу испитивања утицаја начина агрегације оцена појединачних агената у МАС-у на тачност аутоматског оцењивања, у оквиру хипотезе Х4, биће упоређене две варијанте МАС. У првој варијанти, коначна оцена ће се формирати применом једноставних облика агрегације (просека, тежинског просека) оцена које генеришу појединачни агенти. У другој варијанти, у те сврхе биће уведен агент за агрегацију, који ће анализирати оцене појединачних агената и њихових образложења, и на основу њих просудити коначну оцену студентског рада. Поређењем ових варијанти МАС-а над истим скуповима података утврдиће се колики утицај агентски вид агрегације оцена има на тачност аутоматског оцењивања.

За евалуацију тачности свих претходно наведених система, између осталог, користиће се корелација (Pearson) као мера линеарне повезаности са референтним оценама, Mean Absolute Error и Root Mean Squared Error као мере просечног одступања системских од референтних оцена, као и Quadratic Weighted Kappa (QWK) која мери степен сагласности између системских и референтних оцена. Поред квантитативне евалуације, биће спроведена и квалитативна анализа грешака како би се идентификовали разлози и обрасци одступања аутоматски додељених оцена у односу на референтне оцене.

У циљу испитивања утицаја МАС-а на транспарентност и објашњивост процеса аутоматског оцењивања, а у оквиру хипотезе Х5, биће упоређени мултиагентски и моноагентски системи засновани на великим језичким моделима, који генеришу образложења својих одлука у вези са процесом оцењивања. Квалитет генерисаних образложења оцена биће процењен различитим методама у зависности од доступности референтног образложења у скупу података. Код скупова података где није доступно референтно образложење, биће примењен приступ у ком ће независан језички модел евалуирати структуру, утемељеност и аргументованост генерисаних образложења у односу на задатак и

критеријуме оцењивања, применом *LLM-as-a-judge* приступа. Код скупова података где ће бити доступна референтна наставничка образложења оцене, квалитет генерисаних образложења ће бити процењен применом метрика сличности са референтним образложењем, као што су ROUGE и BERTScore.

Сви експерименти биће реализовани у програмском језику Пајтон (Python). За имплементацију MAC-а користиће се библиотеке за велике језичке моделе и агентски оквири (LangChain, LangGraph и сл). За имплементацију приступа спољном знању користиће се библиотеке за рад са векторским базама (FAISS, ChromaDB, Qdrant или сл), као и векторски (embedding) модели доступни путем платформи као што је *HuggingFace*. За обраду података користиће се библиотеке попут *pandas* и *NumPy*. Велики језички модели биће хостовани коришћењем платформе *Ollama* и сл. С обзиром на брз развој нових технологија у области вештачке интелигенције и MAC-а, предложене технологије ће бити подложне промени, те је могуће да ће у току израде дисертације бити коришћени и други алати, библиотеке и платформе које у датом тренутку буду актуелне и погодније за реализацију експеримената, уз задржавање постављене методологије и експерименталног дизајна.

## 7. Мултидисциплинарност теме (максимално 1/2 стране)

Мултидисциплинарни карактер ове теме огледа се у повезивању више научних области. Из угла рачунарских наука, истраживање се бави применом савремених метода вештачке интелигенције, пре свега применом великих језичких модела из области процесирања природних језика, метода развоја MAC-а, као и приступа из области софтверског инжењерства. Истовремено, рад се ослања на педагошке приступе вредновања знања, посебно кроз употребу рубрика за оцењивање у аутоматским системима. Додатно, постоји веза и са едукативном психологијом у домену транспарентности и објашњивости процеса оцењивања, који утичу на поверење наставника и студената у аутоматизоване системе, као и на перцепцију правичности и поузданости процеса аутоматског оцењивања. Повезаност са педагогијом и едукативном психологијом огледа се и кроз примену концепта повратне информације (feedback), који омогућује да аутоматски систем пружи аргументовано образложење додељене оцене студенту и наставнику. Истраживање се може посматрати и у ширем контексту области образовних технологија, јер доприноси развоју интелигентних система намењених подршци процесу наставе и вредновању знања.

## 8. Очекивани научни допринос докторске дисертације

Очекивани научни допринос ове докторске дисертације обухвата дизајн, развој и мултиаспектну евалуацију иновативног приступа објашњивом аутоматском оцењивању студентских радова, заснованом на систему интелигентних агената, који интегрише велике језичке моделе и спољне изворе знања ради унапређења тачности, конзистентности и објашњивости процеса аутоматског оцењивања.

Конкретно, очекивани научни доприноси се огледају у следећем:

- Развој архитектуре MAC-а заснованог на великим језичким моделима за аутоматско оцењивање студентских радова, са приступом спољном знању, који омогућује адаптивну специјализацију агената у односу на тип задатка и по различитим критеријумима оцењивања дефинисаних рубриком за оцењивање, као и механизам за агрегацију појединачних процена агената у јединствену коначну оцену, уз образложење додељене оцене
- Емпиријски утврђен утицај мултиагентског приступа на тачност аутоматског оцењивања у односу на постојеће моноагентске и неагентске приступе, чиме се доприноси разумевању потенцијала примене MAC-а у образовним окружењима
- Емпиријски процењен утицај адаптивне специјализације агената у односу на тип задатка и по појединачним критеријумима оцењивања на тачност аутоматског оцењивања
- Емпиријски процењен утицај различитих начина интеграције спољног знања у MAC на тачност аутоматског оцењивања
- Емпиријски утврђен утицај различитих начина агрегације оцена појединачних агената на коначну оцену MAC-а
- Приступ за транспарентно и објашњиво аутоматско оцењивање кроз генерисање образложења оцене студентског рада, засновано на рубрикама и релевантном спољном знању
- Предлог експерименталног оквира за систематско испитивање MAC-а заснованих на великим језичким моделима, у задацима аутоматског оцењивања студентских радова, који омогућује контролисано поређење различитих варијанти система и конфигурација агената
- Допринос бољем разумевању примене великих језичких модела и мултиагентских архитектура у образовању, као и њихових предности и ограничења у задацима аутоматског оцењивања и пружања повратних информација
- Могућност примене предложеног MAC-а у реалном образовном окружењу, посебно у задатку оцењивања програмског кода

## 9. Закључак и предлог комисије

На основу изнетих података Комисија сматра да је предложена тема докторске дисертације **Мултиагентски приступи засновани на великим језичким моделима за објашњиво аутоматско оцењивање студенских радова** научно заснована и актуелна, и да кандидат Марко Путниковић, мастер информатичар, испуњава услове за рад на овој докторској дисертацији.

Комисија предлаже Већу за студије при Универзитету у Београду да прихвати тему и кандидату Марку Путниковићу одобри израду докторске дисертације под наведеним насловом. За менторе се предлажу др Јелена Јовановић, редовни професор, Факултет организационих наука Универзитета у Београду (ужа научна област: софтверско инжењерство) и др Александар Бауцал, редовни професор, Филозофски факултет Универзитета у Београду (ужа научна област: општа психологија).

У Београду \_\_.3.2026.

Потписи чланова комисије:



Др Ранка Станковић, редовни професор,  
Универзитет у Београду, Рударско-геолошки факултет,  
ужа научна област: математика и информатика



Др Срећко Јоксимовић, професор,  
*Adelaide University, School of Education,*  
ужа научна област: аналитике учења, вештачка  
интелигенција у образовању



Др Владан Девеџић, редовни професор,  
Универзитет у Београду, Факултет организационих наука,  
ужа научна област: софтверско инжењерство