

**УПОТРЕБА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ
ЗА ПРОЦЕНУ РИЗИКА ОД РЕЦИДИВИЗМА
У КРИВИЧНОМ ПОСТУПКУ И ПОШТОВАЊЕ
ПРАВА НА ПРАВИЧНО СУЂЕЊЕ ИЗ ЧЛАНА 6.
СТАВ 1. ЕВРОПСКЕ КОНВЕНЦИЈЕ
О ЉУДСКИМ ПРАВИМА**

Апстракт: У савременим кривично-правним системима користе се сложени модели вештачке интелигенције (ВИ) за процену ризика од рецидивизма, чиме инструменти за превенцију кривичних дела постају алгоритами.

Овај рад анализира појам рецидивизма као облик превенције кривичних дела и инструменте за процену ризика од рецидивизма у кривичном поступку, са посебним освртом на инструменте који користе вештачку интелигенцију и машинско учење као помоћ судијама при доношењу одлука. Такви ВИ модели се већ примењују у САД.

Узимајући у обзир проблеме који су се појавили у пракси применом ових модела у САД, као и опште препознатљиве изазове у употреби вештачке интелигенције, овај рад разматра те проблеме с аспекта поштовања гаранција права на правично суђење из члана 6. став 1. Европске конвенције за заштити људских права и основних слобода. Циљ је утврдити да ли се овакви модели, у свом садашњем облику, могу примењивати у државама-чланицама Савета Европе у складу са прихваћеним стандардима правичности кривичног поступка.

Кључне речи: вештачка интелигенција, процена ризика од рецидивизма, кривични поступак, право на правично суђење, Европска конвенција о људским правима.

* Државно правобранилаштво Републике Србије, Одељење за заступање Републике Србије пред Европским судом за људска права, саветник, jelenasmil@yahoo.com. Ставови које је аутор изнео у раду представљају искључиво личне ставове аутора.

1. УВОД

Употреба вештачке интелигенције¹ (у даљем тексту: ВИ) представља велики изазов данашњице, посебно када се користи у правосуђу, а то се може чинити на два начина. С једне стране, ВИ може помагати судијама приликом доношења судских одлука, док, с друге стране, сама ВИ може узети улогу судија и заменити човека на том месту, самостално доносећи одлуке у судском поступку.

Када је у питању први начин коришћења ВИ у правосуђу, њена улога се своди на помоћ судијама путем предвиђања, односно припреме судских одлука. Такви алгоритми нуде бројна нова и драгоцене решења на различитим пољима, попут превенције кривичних дела, која се може спроводити, између осталог, и путем модела ВИ за процену ризика од рецидивизма.

С друге стране, ове технологије могу имати и негативне последице, посебно на људска права и основне слободе, што је посебно значајно када је у питању њихова употреба у процесу одлучивања у кривичним поступцима.²

2. ПРЕВЕНЦИЈА КРИВИЧНИХ ДЕЛА И ЊЕНИ ТИПОВИ

Превенција кривичних дела сеже више од два века уназад и укључује различите стратегије, чији је циљ смањење стопе криминалитета и заштита друштва у том смислу. Још су либерални филозофи 18. века друштво схватили као систем, којег чине рационални појединци вођени утилитаризмом. Тада се превенција кривичних дела посматрала кроз ту призму, односно промишљеним понашањем смањивали су се ризици од могуће штете. У каснијем периоду, приступ превенцији криминала био је свеобухватнији. Држава и појединци су међусобно сарађивали, стварајући услове у којима је било све теже и ризичније починити кривично дело, све у циљу заштите заједнице. Државе, потом, формирају и полицијске снаге ради превенције криминала. Доносе се нови закони, а финансијска средства се улажу у циљу ефикасније контроле криминала. Настају и бројне стратегије за превенцију криминала, које нису биле само резултат интервенције државе,

1 „Вештачка интелигенција дефинише се као способност рачунарског програма или машине да размишља и учи слично људима, користећи системе вештачких неуронских мрежа, како би јој се спознаја побољшала“, <https://www.sciencedirect.com/topics/computer-science/artificial-intelligence>, 2. март 2025.

2 Guidelines on artificial intelligence and data protection, T-PD(2019)01, Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data (Convention 108), Directorate General of Human Rights and Rule of Law, 2019, 1, <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>, 2. март 2025.

већ и приватног сектора и институција цивилног друштва, што указује на ширење дискурса са државне на друштвену контролу овог питања.³

Крајем 20. века у криминологији се издвојило неколико типова превенције кривичних дела. Законска превенција се спроводи кроз примену закона и има директне и индиректне ефекте, попут одвраћања, онемогућавања и рехабилитације, односно има морално-образовну улогу. Када говоримо о факторима који делују на развој и понашање појединаца, на пољу смо развојне превенције. Развојна превенција се врши тако што се идентификује што већи број ризичних фактора, а потом осмишљавају конкретне превентивне мере којима би се смањили ти ризични фактори. Развојна превенција се често обједињује са превенцијом у заједници, формирајући нови тип, а то је социјална превенција. Она настоји да превенира све факторе који утичу на испољавање криминалног понашања, како на индивидуалном нивоу, тако и на нивоу породице и заједнице. Када се овај елемент реаговања на ране ризичне факторе споји са редуковањем прилике за кривично дело, добија се превенција у заједници. Правац који је доминантан у савременим схватањима превенције јесте ситуациона превенција. Она се не ослања увек на инструменте државне политике и правосуђа, већ то могу бити и, рецимо, приватне безбедносне агенције. Методама ситуационе превенције обухваћен је широк спектар облика криминала, притом је лако применљива и приступачна. Може се спроводити кроз примену прописа, видео надзор, грађанске патроле и слично.⁴

Данас, са развојем нових технологија и употребом ВИ, превенција кривичних дела добија свој нови тип и можемо рећи да постаје алгоритамска. Алгоритми се у кривично-правном систему користе да предвиде где ће највероватније доћи до извршења кривичног дела, као и ко највероватније може први пут или поново извршити кривично дело, као и да се не појави на суђењу.⁵ Један тип ВИ модела за превенцију криминала, који се користи у кривичном поступку, јесте модел за процену ризика од рецидивизма.

2.1. Појам рецидивизма и процена ризика од рецидивизма у кривичном поступку

Рецидивизам представља сложену друштвену појаву за чије разумевање је неопходан целовит приступ, који обухвата правни, социјални и психолошки аспект. У питању је феномен отпоран на бројне мере за његово

3 A. Crawford, K. Evans, „Crime Prevention and Community Safety“, *The Oxford Handbook of Criminology* (eds. A. Leibling, S. Maruna, L. McAra), Oxford University Press, 2017, 797–824.

4 М. М. Ковачевић-Лепојевић, З. Илић, „Ситуациона превенција криминала у контексту“, *Војно дело*, 5/2018, 55–58.

5 J. Dressel, H. Farid, „The accuracy, fairness, and limits of predicting recidivism“, *Science Advances*, 1/2018, 1.

сузбијање и спречавање, а веома прилагодљив друштвеним променама. Појам рецидивизма се дефинише с кривичноправног, криминолошког и пенолошког аспекта. Кривичноправна дефиниција посматра појам рецидивизма с правног становишта, с нагласком на правне последице и санкције за поновно почињено кривично дело, након што је лице осуђено за раније почињено кривично дело. Криминолози настоје да утврде узроке и мотиве који доводе до рецидивизма, узимајући у обзир друштвене, економске и психолошке факторе, који утичу на понашање повратника. С обзиром на то, врше класификацију повратника на деликвенте из навике, деликвенте по тенденцији и професионалне деликвенте. Имајући у виду сложеност и вишеслојност проблема рецидивизма, пенологија се бави ресоцијализацијом и третманом повратника. Наглашава важност индивидуалног приступа, узимајући у обзир посебне, јединствене карактеристике сваког повратника, поред општих друштвених узрока који утичу на понављање криминалног понашања.⁶

Увиђајући значај који процена ризика од рецидивизма може имати на превенцију криминалитета, у кривичним системима настају бројни инструменти за вршење ове процене још од 30-их година прошлог века, који свој врхунац доживљавају пре тридесетак година.⁷ На самом почетку процена ризика се није вршила структурисано и заснивала се на процени медицинских стручњака, односно стручњака из казнено-поправних завода, који нису морали да поштују посебне смернице, већ су процену давали на основу личног професионалног искуства. Развојем технологије, развијају се и начини за процену ризика од рецидивизма и добијају механичке, као и алгоритамске карактеристике. Ризици који се односе на конкретног појединца почињу да се упоређују са референтном групом.⁸

Приликом процене се узимају у обзир карактеристике појединца, које су у директној вези са могућношћу рецидивизма, односно криминогене потребе. Оне се, као фактори ризика, деле на статичке и динамичке. Статички фактори ризика су они које се не могу мењати, попут пола, година и криминалне историје. Они представљају јаке факторе за предвиђање будућег криминалног понашања, али нису подложни било за какву интервенцију када је у питању рехабилитација.⁹ Временом се у ове програме за процену ризика од рецидивизма укључују динамички фактори осетљиви

6 С. Константиновић-Вилић, „Узроци рецидивизма и третман повратника“, *Зборник Правног факултета у Нишу* (ур. Љ. Јовановић), Ниш, 1977, 389–392.

7 M. Ghasemi et al., „The application of the Machine Learning to a General Risk/Need Assessment Instrument in the Prediction of Criminal Recidivism“, *Criminal Justice and Behavior*, 4/2021, 518.

8 E. Chelioudakis, „Risk assessment tools in criminal justice: Is there a need for such tools in Europe and would their use comply with European data protection law?“, *Australian National University Journal of Law and Technology*, 2/2020, 72–80.

9 D. Yukhnenko, Blackwood, S. Fazel, „Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis“, *CNS Spectrums*, 2/2020, 252–263.

на промене, попут неадекватних породичних односа, злоупотребе психоактивних супстанци, радног статуса, образовања и слично. Укључивањем динамичких фактора добија се прецизнија процена ризика, као и свеобухватнији преглед околности које могу утицати на повећање вероватноће да појединац поново изврши кривично дело, притом се на њихову промену може утицати током процеса рехабилитације. С друге стране, динамички фактори ризика предмет су критика, због субјективног карактера, који води непоузданој процени.¹⁰

Данашњи модели за процену ризика од рецидивизма функционишу свеобухватно и систематично, усклађујући лични профил преступника са прилагођеним плановима за рехабилитацију. Рехабилитација преступника се документује кроз факторе који утичу на његово криминално понашање, тако да интервенција може бити усмерена директно на фактор који може довести до смањења рецидивизма. Најновији модели користе стандардизоване факторе за процену ризика, што модел чини објективнијим. Ови модели побољшавају тачност предвиђања техникама машинског учења,¹¹ односно вештачке интелигенције, која подразумева обуку на великом скупу података, што омогућава проналажење трендова, које ранији модели нису могли да уоче.¹²

Иако слични по својој сврси, различити модели користе различите методологије за процену ризика. ОхRec (*Oxford Risk of Recidivism Tool*) за предвиђање насилног рецидивизма код пуштених затвореника користи статистички приступ.¹³ Модел VRAG (*Violence Risk Appraisal Guide*) приликом процене ризика за насилничко понашање комбинује историјске податке и статичке факторе. За потребе идентификације потреба појединаца и ризика који могу утицати на рецидив, са циљем да се развију одговарајући планови интервенције, користи се LSI-R (*Level of Service Inventory-Revised*), док се за процену психопатије, посебно код озбиљних преступника, користи PCL-R (*Psychopathy Checklist-Revised*). Примећује се да називи ових инструмената указују на област процене ризика.¹⁴

10 В. Петровић, Г. Јованић, „Процена ризика од рецидивизма у затворском систему – достигнућа и изазови“, *Безбедносћ*, 2/2022, 27 и 28.

11 „Машинско учење је област вештачке интелигенције која омогућава рачунарима да самостално уче и побољшавају своје перформансе у обављању задатака анализом великих скупова података, без потребе за експлицитним програмирањем сваког корака“, <https://www.merriam-webster.com/dictionary/machine%20learning>, 11. март 2025.

12 E. Chelioudakis, *op. cit.*, 72–80.

13 G. Van Dijk, “Predicting Recidivism Risk Meets AI Act”, *European Journal on Criminal Policy and Research*, 28/2022, 408 и 409.

14 T. Brennan, W. Dieterich, B. Ehret, “Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System”, *Criminal Justice and Behavior*, 1/2009, 22–32; више о моделима за процену ризика од рецидивизма вид. у Г. Јованић, „Кривично-правни и пенолошки рецидивизам осуђених лица“, *Специјална едукација и рехабилитација* (ур. Д. Илић-Стошовић), 2/2010, Београд, 271–282.

2.2. Модели вештачке интелигенције за процену ризика од рецидивизма

У циљу побољшања инструмената, односно модела за процену ризика од рецидивизма, у корак с технолошким развојем, развијали су се и предиктивни модели који користе вештачку интелигенцију и машинско учење. Они се користе у кривичним поступцима појединих правних система као помоћ судијама да донесу одлуку о изрицању казне, условном отпусту, као и рехабилитацији, вршећи процену вероватноће да ли ће оптужени поново извршити кривично дело у скоријем периоду, уколико буде пуштен на слободу. Циљ употребе ових модела није само предвиђање, већ побољшање јавне безбедности и превенција криминала.¹⁵

Један од најпознатијих алгоритама данашњице за предвиђање ризика од рецидивизма је COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), који се користи у америчком кривично-правном систему. На основу криминалне историје појединца, образаца понашања, демографских и социо-економских фактора, нивоа образовања или претходног запослења, овај алгоритам додељује оптуженима оцену ризика од 1 до 10. Притом, оцена 10 указује на највиши степен ризика од рецидивизма.¹⁶ У зависности од квалитета и репрезентативности података којима се похрањују, предвиђања ових програма могу бити или не бити кредибилна. Наиме, наведени подаци се користе за обуку ВИ програма у циљу препознавања фактора ризика за поновно извршење кривичних дела и прављења образаца у том смислу. Функционишу тако што се у програм унесе нови случај, а програм врши анализу карактеристика конкретног појединца и процену вероватноће за рецидивизам.¹⁷ Приликом процене, алгоритам може приписати већу тежину постојању неких прекршаја или кривичних дела из прошлости, али узимајући у обзир и расу којој оптужени припада. Овај податак изазвао је велику забринутост стручне јавности због упитне правичности целог поступка.¹⁸ У стручној јавности постоји став да ће алгоритми који користе машинско учење за обуку на подацима, који генеришу људску пристрасност, давати, а не елиминисати такође пристрасне резултате.¹⁹

Алгоритам COMPAS се примењује у САД и веома су подељена мишљења око његове примене у кривично-правном систему те државе. Док неки сматрају да се његовом применом може реформисати кривично правосуђе

15 M. Ghasemi *et al.*, *op. cit.*, 518.

16 D. Purves, J. Davis, "Should Algorithms that Predict Recidivism Have Access to Race?", *American Philosophical Quarterly*, 2/2023, 205–220.

17 M. M. Farayola *et al.*, "Ethics and Trustworthiness of AI for Predicting the Risk of Recidivism: A Systemic Literature Review", *Information*, 8/2023, 1.

18 D. Purves, J. Davis, *op. cit.*, 205–220.

19 K. Lum, "Limitations of mitigating judicial bias with machine learning", *Nature Human Behaviour* 1(7)/2017.

у САД, које је оптерећено расним проблемима већ дуже време, други сматрају да употреба овог система управо подстиче расни проблем и води ка дискриминацији.²⁰

У правосудним системима европских држава не користе се модели ВИ за процену ризика од рецидивизма, али су се у појединим земљама вршила одређена истраживања и спроводили пројекти. Према извештају СЕРЕЈ (*Council of Europe's European Commission for the Efficiency of Justice*), у неким европским земљама, попут Велике Британије, спроводили су се пилот програми за евентуално коришћење ВИ у процени ризика од рецидивизма, односно у Француској, Холандији и Грчкој за коришћење у правосудју генерално.²¹ Такође, неки програми машинског учења за процену ризика од рецидивизма истраживали су се у Шведској и Холандији.²²

Ни у Републици Србији се не користе предиктивни ВИ модели овог типа. Наиме, иако је процена ризика од рецидивизма након 2013. године са неуједначених критеријума прешла на примену инструмената процене, и даље се не примењују модели машинског учења, односно вештачке интелигенције.²³

Европска Унија је 2024. године прва на свету донела Уредбу о вештачкој интелигенцији, која је ступила на снагу 1. августа 2024. године. Уредбом се настоји да се регулише овај велики глобални технолошки изазов и хармонизују правила о хуманоцентричној примени ВИ, штитећи Унију од могућих штетних последица. Овај акт идентификује степен ризика од примене ВИ система по европске вредности, међу којима су и људска права и право на правично суђење. Наиме, што је опасност по друштво већа, то је и регулатива за примену конкретних програма ВИ строжа, а ризик се креће од минималног до високог, уз постојање и оних ВИ програма који су забрањени. Притом, Акт идентификује бројна поља примене ВИ, међу којима је и правосудје.²⁴

Анализом овог правног акта, може се закључити да је применљив и на ВИ моделе за процену ризика од рецидивизма. Уредба их сврстава у забрањене ВИ системе, али их не забрањује апсолутно, већ условно. Наиме, физичком лицу се може судити једино на основу његовог стварног понашања, а не на основу предвиђања ВИ која се заснивају искључиво на профилисању особа и личним карактеристикама и особинама, попут националности, места рођења и пребивалишта, брачног, друштвеног и финансијског статуса,

20 E. Chelioudakis, *op. cit.*, 72 и 73.

21 E. Chelioudakis, *op. cit.*, 73 и 74.

22 G. Van Dijk, *op. cit.*, 408 и 409.

23 В. Петровић, Г. Јованић, *op. cit.*, 29 и 30.

24 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 – Artificial Intelligence Act, *Official Journal of the European Union* L series, 12. 7. 2024, (1), (4), (48).

а без узимања учешћа људске процене. Дакле, уколико би ВИ систем самостално вршио процену ризика од рецидивизма, а заснивао је само на профилисању особе и процени њених личних карактеристика и особина, примена ових модела ВИ била би забрањена. Ова одредба почела је да се примењује од 2. фебруара 2025. године.²⁵ Међутим, уколико ВИ систем не заснива своју процену ризика на претходно наведеном, већ на објективним и проверљивим чињеницама, а као такав служи као помоћни алат у спровођењу закона, уз постојање људске процене, ВИ модел за процену ризика од рецидивизма може се користити у кривичним поступцима.²⁶ Иако би тада примена ових алгоритама била дозвољена, спадала би у високо ризичне,²⁷ тако да би ови алгоритми морали да испуне строге регулаторне услове, пре него се почне са њиховом применом, а чији је рок за имплементацију 2. август 2026. године.²⁸

Имајући у виду да се под одређеним условима примена модела ВИ за процену ризика од рецидивизма дозвољава у државама-чланицама Европске Уније, да се у осталим државама-чланицама Савета Европе не забрањује, а да су све ове државе потписнице Европске конвенције о заштити људских права и основних слобода (у даљем тексту: Конвенција), у наставку ћемо анализирати гаранције права на правично суђење из члана 6, став 1 Конвенције и тумачење овог члана у пракси Европског суда за људска права (у даљем тексту: Суд), а потом и могуће повреде гаранција уколико се у правне системе европских земаља укључи примена савремених ВИ модела за процену ризика од рецидивизма.

3. ГАРАНЦИЈЕ ПРАВА НА ПРАВИЧНО СУЂЕЊЕ У КРИВИЧНИМ ПОСТУПЦИМА ИЗ ЧЛ. 6, СТ. 1 ЕВРОПСКЕ КОНВЕНЦИЈЕ О ЉУДСКИМ ПРАВИМА И ЊИХОВА МОГУЋА ПОВРЕДА УПОТРЕБОМ ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ ЗА ПРОЦЕНУ РИЗИКА ОД РЕЦИДИВИЗМА

Да би се програми вештачке интелигенције могли интегрисати у правосудни систем, неопходно је да испуњавају стандарде правичног суђења.²⁹ Када говоримо о државама-чланицама Савета Европе, међу којима је и Република Србија, њихов правни систем мора бити усклађен са Европском конвенцијом, као и праксом Европског суда. Када је у питању право на

25 Чл. 5 (д), Поглавље 2; чл. 113. Поглавље 13, Artificial Intelligence Act.

26 Тачка (42); (59); чл. 5 (д), Поглавље 2, Artificial Intelligence Act.

27 Анекс 3. тачка 6 (д), Artificial Intelligence Act.

28 Чл. 8, Секција 2, Поглавље 3; чл. 113, Поглавље 13, Artificial Intelligence Act.

29 K. Terzidou, "The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial", *Journal of Judicial Administration*, 3/ 2022, 158.

правично суђење, релевантне гаранције су прописане у члану 6 Конвенције, а ми ћемо се фокусирати на оне које су прописане у ставу 1 овога члана.

Гаранције права на правично суђење из члана 6, став 1 Конвенције односе се, између осталог, и на кривичне поступке, при чему се кривични поступак посматра као целина, укључујући и истражни поступак, односно фазу претходног поступка.³⁰ Гаранције права на правично суђење изричито прописане овом одредбом Конвенције су: правична и јавна расправа, са изузецима, независан и непристрасан законито основан суд, као и суђење у разумном року.³¹ Међутим, тумачећи ову одредбу, Суд је током времена кроз своју праксу развио још неке гаранције, попут права на приступ суду, као и права на равноправност и акузаторност поступка.³² Да би Суд констатовао да постоји повреда права на правично суђење, довољно је да буде повређена само једна од наведених гаранција.

Говорећи генерално о употреби вештачке интелигенције у судским поступцима, она има за циљ да побољша процес доношења одлука, поступке учини ефикаснијим и судску праксу уједначенијом. ВИ се у том смислу користи да судијама пружи аналитичку подршку и помогне им у њиховом дискреционом резоновању.³³ И модели ВИ за процену ризика од рецидивизма, такође, могу учинити кривични поступак ефикаснијим, али и учинити судску праксу уједначенијом, с обзиром на то да омогућавају стандардизацију процене ризика за рецидивизам и доследно третирање сличних случајева. Међутим, технолошки напредак не сме бити на штету по поштовање људских права, тако да, иако постоје одређене користи од примене ВИ модела, уколико постоји повреда само неке од гаранција права на правично суђење у домаћим кривичним поступцима у којима се ови модели буду користили у државама-потписницама Конвенције, ризик по људска права превазићи ће корист од њихове примене.

Иако у пракси Европског суда постоје случајеви који се односе на преиначење кривичне санкције у којима је Суд као критеријум за доношење одлуке узимао и ризик од рецидивизма,³⁴ као и случајеви који се тичу употребе алгоритама у кривичним поступцима,³⁵ до сада се пред Судом није

30 *Salduz v. Turkey*, бр. представке 36391/02, ст. 50, пресуда ВВ ЕСЉП од 26. новембра 2008; *Vera Fernández-Huidobro v. Spain*, бр. представке 74181/01, ст. 111, пресуда ЕСЉП од 6. јануара 2010.

31 Чл. 6, ст. 1, Европске конвенције за заштиту људских права и основних слобода, *Службени лист СЦГ – Међународни уговори*, бр. 9/03, 5/05 и 7/05 – испр. и *Службени гласник РС – Међународни уговори*, бр. 12/10.

32 Ј. С. Миленковић, *Право на правично суђење у грађанским споровима у пракси Европског суда за људска права*, Службени гласник, ФЕППС, 2019, 103 и 147.

33 G. Sartor, K. L. Branting, "Introduction: Judicial Applications of Artificial Intelligence", *Judicial Applications of Artificial Intelligence* (eds. G. Sartor, K. Branting), Springer, Dordrecht, 1998, 105–109.

34 *Inter alia Dardanskis and Others v. Lithuania*, бр. представке 74452/13, одлука ЕСЉП од 18. јуна 2019.

35 *Sigurður Einarsson and Others v. Iceland*, бр. представке 39757/15, пресуда ЕСЉП од 4. јуна 2019.

појавио случај употребе модела ВИ за процену ризика од рецидивизма у домаћим кривичним поступцима држава-чланица Савета Европе, односно потписница Конвенције. Стога ћемо анализом појединих гаранција правичности кривичног поступка прописаних чл. 6, ст. 1 Конвенције и дефинисаних у пракси Европског суда, с једне стране, као и анализом недостатака модела ВИ за процену ризика од рецидивизма, с друге стране, пробати да дођемо до закључка како би и под којим условима Суд одлучио, уколико би употреба једног таквог програма ВИ била спорна пред Судом у смислу поштовања права на правично суђење.

3.1. Могућа повреда права на независан и непристрасан суд из чл. 6, ст. 1 Конвенције коришћењем вештачке интелигенције за процену ризика од рецидивизма у кривичном поступку

Да би се кривични поступак сматрао правичним, неопходно је да се води пред независним и непристрасним законом основаним судом. Иако у правној теорији постоји јасна разлика између појмова независности и непристрасности, Суд констатује да их је тешко раздвојити, јер их подносиоци у својим представкама истовремено доводе у питање и заснивају повреду на истим чињеницама. Стога, Суд истовремено преиспитује повреду ових гаранција.³⁶

Начелно говорећи, независност представља слободу правосуђа и судија од спољашњих и унутрашњих непримерених утицаја, док се непристрасност посматра кроз однос судија према странкама и према самом случају у коме поступају.³⁷

Притисци и недостојни утицаји на судску власт у пракси обично долазе од стране извршне и законодавне власти кроз именовање судија, трајање судијских мандата, непоштовање принципа случајног судије и финансирање судија.³⁸ Међутим, непримерен утицај на поступајућег судију може доћи и изнутра, и то од стране његових колега или судске управе, када ће независност суда подједнако бити повређена.³⁹

36 J. С. Миленковић, *op. cit.*, 71.

37 Више о праву на независан и непристрасан суд у пракси Европског суда за људска права вид. у *ibid.*, 71–92.

38 *Moiseyev v. Russia*, бр. представке 62936/00, ст. 175–185, пресуда ЕСЉП од 9. октобра 2008; *Miroshnik v. Ukraine*, бр. представке 75804/01, ст. 63, пресуда ЕСЉП од 27. новембра 2008; *Oleksandr Volkov v. Ukraine*, бр. представке 21722/11, ст. 130, пресуда ЕСЉП од 27. маја 2013; *Luka v. Romania*, бр. представке 34197/02, ст. 37–50, пресуда ЕСЉП од 21. јула 2009; *Henryk Urban and Ryszard Urban v. Poland*, бр. представке 23614/08, ст. 53–54, пресуда ЕСЉП од 28. фебруара 2011; J. С. Миленковић, *op. cit.*, 81 и 82.

39 Judge G. Borkowski, “Judicial Independence in the Light of art. 6. of the European Convention of Human Rights- selected aspects”, *Teka Komisji Prawniczej – OL PAN*, 2014,

Да би постојала повреда права на независан суд, није нужно доказати непримерени утицај, већ је довољно да постоји утисак да је ова гаранција правичности кривичног поступка угрожена. Притом, сумње се морају објективно оправдати, посматрано из угла не само оптуженог, већ и „објективног посматрача“, који би имао разлога да дели исту забринутост за независност суда у околностима конкретног случаја.⁴⁰ Такође, утицај на суд може да постоји, али, све док судија не подлегне том утицају, његово поступање ће се сматрати независним и непристрасним.⁴¹

Тумачећи појам непристрасности, Суд ову гаранцију посматра са два аспекта. Првенствено, не смеју постојати предрасуде или пристрасности у субјективном смислу. С друге стране, суд мора бити и објективно непристрасан, односно мора постојати легитимна сумња у његову непристрасност.⁴² Дакле, Суд је формулисао тест за оцену непристрасности поступајућег органа кроз субјективни и објективни приступ, а њихова примена ће зависити од чињеница сваког конкретног случаја који се пред њим нашао. И сам Суд је констатовао да је веома тешко установити повреду на основу субјективног теста, те се најчешће фокусира на тест објективне непристрасности. Тада Суд разматра да ли постоје чињенице које могу да изазову сумњу у непристрасно поступање судије, а да се те чињенице могу и проверити.⁴³

Од модела вештачке интелигенције за процену ризика од рецидивизма очекивало се да буде објективнији и праведнији у односу на људско просуђивање. Управо због овог предубеђења, коришћење модела доводи до лажног осећаја сигурности у његову правичност, како због могућег утицаја на независност судског поступка, тако и због тога што се у пракси појављује озбиљна забринутост због његове пристрасности. Наиме, овај проблем није само техничке природе, већ се манифестовао у погледу статистике, етике и права.⁴⁴

Ако бисмо поставили питање независности кривичног поступка у коме се примењују ВИ модели, требало би поћи од тога да ове моделе не развија држава, иако их држава може финансирати, већ приватне компаније. Самим тим, оне утичу не само на функцију и дизајн, већ се кроз ове моделе могу фаворизовати неки интереси, који су свакако финансијски, али могу бити и политички, чиме би се могла нарушити независност суда који користи овакве ВИ моделе. Додатни проблем је што се овакав непримерени утицај

12–13; *Daktaras v Lithuania*, бр. представке 42095/98, ст. 36, пресуда ЕСЉП од 10. октобра 2000.

40 *Incal v. Turkey*, бр. представке 22678/93, пресуда ЕСЉП од 9. јуна 1998, ст. 71; *Clarke v. the UK*, бр. представке 23695/02, одлука ЕСЉП од 25. августа 2005.

41 J. С. Миленковић, *op. cit.*, 76 и 77.

42 *Morice v. France*, бр. представке 29369/10, ст. 73, пресуда ЕСЉП од 23. априла 2015; *Clarke v. the UK*.

43 *Kyprianou v. Cyprus*, бр. представке 73797/01, ст. 119–121, пресуда ЕСЉП од 15. децембра 2005.

44 D. Danks, A. J. London, "Algorithmic Bias in Autonomous Systems", *International Joint Conference on Artificial Intelligence*, 2017, 4691.

не може истражити и контролисати. С друге стране, уколико би се ови програми правили од стране државе, ту би свакако морала бити ангажована извршна власт, пошто има одговарајући стручни кадар за реализацију таквих програма, па би се овакви утицаји извршне власти на судску могли сматрати непримереним.⁴⁵ Имајући у виду наведену праксу Суда, могли бисмо констатовати да би примена оваквих ВИ модела представљала повреду права на независан суд.

Што се тиче проблема алгоритамске пристрасности ВИ модела за процену ризика од рецидивизма, узрок може лежати у подацима на којима се алгоритми обучавају, имајући у виду да се подаци заснивају на ранијим кривичним евиденцијама. Рецимо, евиденције могу указивати на прекомеран криминалитет у одређеним заједницама, попут урбаних средина или заједница већински настањених припадницима одређене расе, попут Афро-Американаца. Учећи на тим подацима, алгоритам ће повезати припаднике црне расе са већом вероватноћом од рецидивизма, иако немају других фактора који указују на то да постоје већи ризици за поновно извршење кривичног дела. Дакле, ово алгоритамско предвиђање не заснива се на стварном понашању појединца чији се ризици за рецидивизам процењују, већ на пристрасним историјским подацима. Самим тим ће алгоритми одступати од етичких стандарда и моралних норми, означавајући појединце ризичним за рецидивизам само на основу расе или друштвено-економског статуса, односно фактора који се етички не могу оправдати.⁴⁶

Конкретно говорећи, у пракси се показало да COMPAS програм за процену ризика од рецидивизма у свом предвиђању показује висок степен расне дискриминације према црнцима. Наиме, чак је двоструко лажно позитивних процена овог програма које се односе на црнце у односу на белу популацију учинилаца кривичних дела. Чак се примећује и сексизам у проценама ризика од рецидивизма, имајући у виду то да у пракси жене показују нижу стопу рецидивизма у односу на мушкарце, што се не може приметити у предвиђањима овог програма.⁴⁷

Такође, нису сви фактори подједнако важни за сваку популацију. Уколико их алгоритам не одмерава на одговарајући начин, резултат предвиђања ризика биће неправичан.⁴⁸

Уколико би се такав алгоритам за процену ризика од рецидивизма користио у кривичним поступцима држава-чланица Савета Европе, помажући судијама да одлуче у конкретном случају, с обзиром на проблем са пристрасношћу алгоритма, сам поступајући суд се не би могао сматрати независним и непристрасним, имајући у виду горе поменућу праксу Европског суда. С обзиром на то да има озбиљних назнака да ови алгоритми у

45 K. Terzidou, *op. cit.*, 160 и 161.

46 D. Danks, A. J. London, *op. cit.*, 4691 и 4692.

47 M. M. Farayola, *op. cit.*, 2.

48 D. Danks, A. J. London, *op. cit.*, 4693.

високом степену показују дискриминацију, уколико би судија донео одлуку на основу предвиђања оваквог програма, постојала би легитимна сумња у независност и непристрасност поступајућег суда, а Европски суд би констатовао повреду права на правично суђење. Осим што би конкретан кривични поступак био неправичан, могао би се довести у питање и цео систем кривичног правосуђа и поверење како оптужених, тако и целокупне јавности, услед коришћења оваквог програма за потребе кривичних поступака.

3.2. Могућа повреда права на образложену судску одлуку из чл. 6, ст. 1 Конвенције коришћењем вештачке интелигенције за процену ризика од рецидивизма у кривичном поступку

Да би се правда могла спровести, Суд у својој пракси констатује обавезу домаћег суда да адекватно образложи разлоге на којима заснива своју одлуку у кривичном поступку. То конкретно значи да стране у поступку морају добити детаљно објашњење сваког аргумента, који је био одлучујући за исход конкретног поступка, и то на начин да одлуку могу разумети.⁴⁹ Због тога је нужно да судска одлука буде образложена јасно и разумљиво, како због оптуженог, тако и због упознавања јавности са њеном садржином путем објављивања. Притом, тврдње у одлуци морају бити објективне. На овај начин се постиже транспарентност судског поступка, доприноси развоју судске праксе и правне доктрине, поверењу грађана у правосуђе, као и заштити права на правично суђење.

Без одговарајуће образложене судске одлуке, није могуће ни оспоравати одлуку, односно није могуће остварити принцип двостепености поступка.⁵⁰ Наиме, када странке разумеју разлоге за доношење конкретне одлуке, могу проценити да ли има основа за њено побијање, подношењем правног лека против те одлуке.

Конечно, поред процедуралних правила и двостепености поступка, и правна сигурност се штити обавезом судије да образложи своју правну логику, односно разлоге на основу којих је дошао до конкретног закључка, а то укључује утврђивање чињеница, њихову анализу и правну оцену.⁵¹

Предиктивни ВИ модели користе машинско учење, које функционише аутономно приликом доношења одлука. Додатној сложености ових модела

49 *Morreira Ferreira v. Portugal* (no. 2), бр. представке 19867/12, ст. 84, пресуда ВВ ЕСЉП од 11. јула 2017.

50 М. Dymitruk, "The Right to a Fair Trial in Automated Civil Proceedings", *Masaryk University Journal of Law and Technology*, 1/2019, 37 и 38.

51 Н. Surden, "The Ethics of Artificial Intelligence in Law: Basic Questions", *The Oxford Handbook of Ethics of AI* (eds. М. D. Dubber, F. Pasquale, S. Das), 2020, 732.

доприноси коришћење техника дубоког учења,⁵² која се налази у сржи најсавременијих програма машинског учења, која користи сложене неуралне мреже⁵³ тешке за разумевање. Један од ВИ модела, код кога се користе ове сложене технике јесте и модел за процену ризика од рецидивизма. Постоје једноставнији ВИ модели који пружају јасне увиде у начин на који су донесене одлуке, али нису у могућности да се изборе са сложеним задацима попут процене ризика од рецидивизма.⁵⁴ Дакле, постизање тачности предвиђања иде на рачун транспарентности, разумљивости и објашњивости програма, који је иначе намењен за коришћење у кривичном поступку.

Иако концепти транспарентности, разумљивости и објашњивости на први поглед делују веома слично, међу њима постоји разлика, иако дефиницијама ових појмова недостаје математичка формалност.⁵⁵

Да би неки програм био транспарентан, потребно је да његов код, подаци на којима се обучава, начин рада и одлучивања буду доступни и јасни. У супротном, настаје феномен „црне кутије“.⁵⁶ Објашњивост се односи на сам процес унутар система машинског учења, односно на могућност да се на јасан начин објасни и опише како је ВИ систем дошао до неке одлуке. У литератури се често транспарентност поистовећује са објашњивошћу, те ћемо и ми за потребе овог рада ова два појма тако посматрати. За разлику од тога, разумљивост, односно интерпретабилност подразумева да човек, конкретно судија или оптужени, може да разуме оно што програм ради, његову логику иза одлучивања и узрочно-последичну везу између података које програм користи у свом предвиђању и одлуке коју је донео. Дакле, разумевање неког програма од стране људи зависи од степена његове транспарентности, односно објашњивости.⁵⁷

Европска комисија је 2019. године објавила Етичке смернице за ВИ од поверења у којима констатује да је за правичност судског поступка неопходно да ВИ сваку своју одлуку објасни људима на разумљив начин,

52 „У основи, модели дубоког учења су регресори података који предвиђају континуирану променљиву на основу скупа улазних података“, <https://www.sciencedirect.com/topics/chemical-engineering/deep-learning>, 10. април 2025.

53 „Неурална мрежа је систем за учење који захтева почетни поступак обуке ради прилагођавања тежинских коефицијената, како би се побољшале перформансе мреже“, <https://www.sciencedirect.com/topics/engineering/learning-systems>, 10. април 2025.

54 P. Linardatos, V. Papastefanopoulos, S. Kotsiantis, “Explainable AI: A Review of Machine Learning Interpretability Methods”, *Entropy*, 1/2021, 1 и 2.

55 *Ibidem*.

56 „Црна кутија“ представља „било шта што има мистериозни или непознати унутрашњи начин рада или механизам“, <https://www.merriam-webster.com/dictionary/black%20box>, 11. април 2025.

57 N. Balasubramaniam et al., “Transparency and Explainability of AI Systems: Ethical Guidelines in Practice”, *Requirements Engineering: Foundation for Software Quality* (eds. V. Gervasi, A. Vogelsang), Springer, 2022, 3–18.

као и да се може испратити њен поступак одлучивања.⁵⁸ Дакле, када говоримо о теми овог рада, закључује се да правичност кривичног поступка зависи од разумљивости и транспарентности, односно објашњивости предиктивног ВИ модела за процену ризика од рецидивизма, уколико се користи у поступку.

Видели смо да и Европска конвенција у гаранције правичног суђења сврстава постојање образложене судске одлуке.

Проблем са ВИ алгоритмима је што се приликом њихове израде приоритет не ставља често на ове важне поменуте факторе, како због сложености алгоритама, недостатка стандардизације, заштите приватности података на којима се тренира алгоритам, посебно оних осетљивих и личних, као и због заштите права интелектуалне својине.⁵⁹

Проблем са транспарентношћу, односно објашњивошћу сложених ВИ системима, посебно оних који се заснивају на дубоком учењу и машинском учењу, попут програма за процену ризика од рецидивизма, је у томе што се често не може објаснити којим процесом је ВИ дошла до одређених резултата, односно коју комбинацију података је користила, да би дошла до конкретног закључка. У питању је такозвани феномен „црне кутије“. Иако су у старту програмери упознати са конкретним алгоритамским кодом, током машинског учења ови алгоритми у ходу мењају првобитни изворни код. С друге стране, оригинални код и алгоритамски процес представљају интелектуалну својину, тако да не могу ни бити откривени кључни елементи програма на којима се суштински заснива предвиђање.⁶⁰ Стога се одлука предиктивног модела не може смислено образложити странкама на које утиче. Притом се угрожава и двостепеност поступка јер се не могу утврдити повреде у поступку, с обзиром на то да је образложење нејасно.⁶¹

Што се тиче разумљивости, као мере у којој људи могу конзистентно да предвиде резултате неког модела и разумеју зашто је донесена баш та конкретна одлука,⁶² последично са недостатком транспарентности, односно објашњивости долази до тога да судије у кривичном поступку погрешно протумаче резултате модела, имајући у виду да нису добро објашњени. Самим тим, судска одлука, заснована на алгоритамском предвиђању, неће бити одговарајуће образложена.

58 *Ethics guidelines for trustworthy AI*, European Commission: Directorate-General for Communications Networks, Publications Office 20, 2019, <https://data.europa.eu/doi/10.2759/34672020>, 12. април 2025.

59 https://www.cimplifi.com/resources/transparency-explainability-and-interpretability-of-ai/?utm_.

60 K. Blount, "Seeking Compatibility in Preventing Crime with Artificial Intelligence", *Masaryk University Journal of Law and Technology*, 1/2021, 42–43.

61 J. Gryz, M. Rojszczak, "Black box algorithms and the rights of individuals: no easy solution to the 'explainability' problem", *Internet Policy Review*, 2/2021, 6–11.

62 H. Surden, *op. cit.*, 731.

3.3. Могућа повреда права на равноправност и акузаторност из чл. 6, ст. 1 Конвенције коришћењем вештачке интелигенције за процену ризика од рецидивизма у кривичном поступку

Правичност поступка не може се замислити без принципа равноправности. Ова гаранција је у уској вези са забраном дискриминације и подразумева једнаку процесноправну заштиту за све стране у поступку, без обзира на различитости било по ком основу, односно представља могућност добијања разумне прилике да свака страна изнесе своје аргументе.⁶³

У једном од случајева, у коме је Суд утврдио да је повређена равноправност странака у поступку, оптуженом је ограничен увид у списе предмета и остале документе. На овај начин оптуженом се ускрађује могућност да оствари увид у доказе, као и прилика да поднесе своје примедбе, чиме се истовремено повређује и право на акузаторни поступак. Наиме, поступајући орган је дужан да одбрани изнесе све материјалне доказе, који су од директног значаја за чињенично стање предмета, посебно ако се суд на те доказе позивао, како би установио кривицу оптуженог, као и ако су ти докази могли да утичу на смањивање казне или ослобађање оптуженог. Уколико докази никада нису откривени, а случај се нашао пред Европским судом, с обзиром на то да Суд не би могао да изврши увид у те доказе, Суд ће морати пажљиво да размотри поступак доношења домаће одлуке, као и да утврди да ли су постојале мере за заштиту интереса оптуженог. До додатне сложености долази уколико се докази односе на електронске податке. У том случају је потребно обезбедити одбрани да се укључи у утврђивање критеријума на основу којих се вршило филтрирање релевантних података.⁶⁴

Да не би дошло до повреде права на акузаторни поступак, није довољно само да докази буду доступни, већ и приступачни, што посебно може бити проблем приликом коришћења напредних технологија. У једном случају пред Судом (*Sigurdur Einarsson a. o. v. Iceland*), докази против подносиоца представке нису били приступачни, иако јесу били доступни, с обзиром на изузетну сложеност дигиталног система коришћеног за прикупљање доказа. Самим тим, одбрана није располагала одговарајућим ресурсима да се припреми, што је довело до повреде права на правично суђење.⁶⁵

Констатовали смо да стране у кривичном поступку морају имати могућност да изнесу своје аргументе и оспоре доказе и анализе које суд користи, на чему се заснива принцип равноправности и акузаторности поступка. Међутим, уколико се у кривичним поступцима користи модел ВИ за процену ризика од рецидивизма, његов проблем са транспарентношћу, односно

63 J. С. Миленковић, *op. cit.*, 147.

64 Водич за члан 6 Конвенције – Право на правично суђење (кривични аспект), 34–38, https://ks.echr.coe.int/documents/d/echr-ks/guide_art_6_criminal_srp, 11. април 2025.

65 K. Blount, *op. cit.*, 42.

објашњивошћу и разумљивошћу, о коме се говорило подробније у претходном делу, може угрозити право оптуженог да оспори доказе или анализе овог алгорита, на основу којих је оптужени оцењен као високо ризичан да поново изврши кривично дело, а који ће бити сутерисани судији да их узме у обзир приликом доношења одлуке.

Феномен „црне кутије“ је и у овом случају присутан, који указује на то да је процес доношења одлука од стране овог ВИ модела нетранспарентан. Наиме, оптужени не може имати приступ методологији или свим подацима на основу којих је процена ризика у његовом случају извршена. Самим тим, ако се оптуженом не оставља могућност да разуме како је процена ризика извршена, он неће моћи ни да оспори конкретне налазе алгорита. У том случају долази до повреде права на равноправан и акузаторни поступак.

Такође, проблем с пристрасношћу ових модела по питању расних и друштвено-економских разлика води неправедном поступању према одређеним групама, самим тим и до неједнаког поступања према странама у поступку.

4. ЗАКЉУЧАК

Питање је времена када ће предиктивни модели вештачке интелигенције, међу којима и модели за процену ризика од рецидивизма, почети да се користе у државама-чланицама Савета Европе и Републици Србији, као део кривичног поступка.

У Етичким смерницама за ВИ од поверења се констатује да је за повећење људи у ВИ неопходно да ВИ поступа етично, безбедно, засновано на подацима, транспарентно, без дискриминације, правично, одговорно, поштујући људску аутономију, приватност, друштвено и еколошко благостање.⁶⁶ Уколико само један елемент недостаје, ВИ може проузроковати већу штету по друштво, него корист. Имајући у виду да је један од кључних елемената за кредибилност ВИ модела поштовање принципа правичности, а да смо у претходном тексту констатовали да предиктивни ВИ модели за процену ризика од рецидивизма не би могли да испуњавају стандард правичног суђења према чл. 6, ст. 1 Конвенције, употреба ових алгорита би нашкодила правима и интересима како појединаца, тако и целог друштва, бар до тренутка док се не реше проблеми пристрасности, транспарентности, односно објашњивости и разумљивости ових алгорита.

Стога, да би се могло говорити о увођењу оваквих алгоритамих система у европско и домаће кривично правосуђе, неопходно је осигурати првенствено широку дебату стручне и академске јавности, јер би примена ових предиктивних ВИ модела имала утицај на целокупно друштво, не само на појединце на чија се права односи.

ЛИТЕРАТУРА

- Balasubramaniam N. et al., "Transparency and Explainability of AI Systems: Ethical Guidelines in Practice", *Requirements Engineering: Foundation for Software Quality* (eds. V. Gervasi, A. Vogelsang), Springer, 2022.
- Blount K., "Seeking Compatibility in Preventing Crime with Artificial Intelligence", *Masaryk University Journal of Law and Technology*, 1/2021.
- Borkowski G., "Judicial Independence in the Light of art. 6. of the European Convention of Human Rights – selected aspects", *Teka Komisji Prawniczej – OL PAN*, 2014.
- Brennan T., Dieterich W., Ehret B., "Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System", *Criminal Justice and Behavior*, 1/2009.
- Van Dijck G., "Predicting Recidivism Risk Meets AI Act", *European Journal on Criminal Policy and Research* 28/2022.
- Водич за члан 6 Конвенције – Право на правично суђење (кривични аспекти), https://ks.echr.coe.int/documents/d/echr-ks/guide_art_6_criminal_srp.
- Ghasemi M. et al., "The application of the Machine Learning to a General Risk/Need Assessment Instrument in the Prediction of Criminal Recidivism", *Criminal Justice and Behavior*, 4/2021.
- Gryz J., Rojszczak M., "Black box algorithms and the rights of individuals: no easy solution to the 'explainability' problem", *Internet Policy Review*, 2/2021.
- Danks D., London A. J., *Algorithmic Bias in Autonomous Systems*, *International Joint Conference on Artificial Intelligence*, 2017.
- Dressel J., Farid H., "The accuracy, fairness, and limits of predicting recidivism", *Science Advances*, 1/2018.
- Dymitruk M., "The Right to a Fair Trial in Automated Civil Proceedings", *Masaryk University Journal of Law and Technology*, 1/2019.
- Јованић Г., „Кривично-правни и пенолошки рецидивизам осуђених лица“, *Специјална едукација и рехабилитација* (ур. Д. Илић-Стошовић), 2/2010, Београд.
- Chelioudakis E., "Risk assessment tools in criminal justice: Is there a need for such tools in Europe and would their use comply with European data protection law?", *Australian National University Journal of Law and Technology*, 2/2020.
- Константиновић-Вилић С., „Узроци рецидивизма и третман повратника“, *Зборник Правног факултета у Нишу* (ур. Љ. Јовановић), Ниш, 1977.
- Ковачевић-Лепојевић М. М, Илић З., „Ситуациона превенција криминала у контексту“, *Војно дело*, 5/2018.
- Crawford A., Evans K., „Crime Prevention and Community Safety“, *The Oxford Handbook of Criminology* (eds. A. Leibling, S. Maruna, L. McAra), Oxford University Press, 2017.
- Linardatos P., Papastefanopoulos V., Kotsiantis S., "Explainable AI: A Review of Machine Learning Interpretability Methods", *Entropy*, 1/2021.

- Lum K., "Limitations of mitigating judicial bias with machine learning", *Nature Human Behaviour*, 1(7), 1, 2017.
- Миленковић Ј. С., *Право на правично суђење у ираћанским ојоровима у пракси Евројској суда за људска права*, Службени гласник, ФЕППС, 2019.
- Петровић В., Јованић Г., „Процена ризика од рецидивизма у затворском систему – достигнућа и изазови“, *Безбедност*, 2/2022.
- Purves D., Davis J., "Should Algorithms that Predict Recidivism Have Access to Race?", *American Philosophical Quarterly*, 2/2023.
- Sartor G., Branting K. L., "Introduction: Judicial Applications of Artificial Intelligence", *Judicial Applications of Artificial Intelligence* (eds. G. Sartor, K. Branting), Springer, Dordrecht, 1998.
- Surden H., "The Ethics of Artificial Intelligence in Law: Basic Questions", *The Oxford Handbook of Ethics of AI* (eds. M. D. Dubber, F. Pasquale, S. Das), 2020.
- Terzidou K., "The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial", *Journal of Judicial Administration*, 3/2022.
- Yukhnenko D., Blackwood, Fazel S., "Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis", *CNS Spectrums*, 2/2020.
- Farayola M. M. et. AL., "Ethics and Trustworthiness of AI for Predicting the Risk of Recidivism: A Systemic Literature Review", *Information*, 8/2023.

ПРАВНИ ИЗВОРИ

- Ethics guidelines for trustworthy AI, European Commission: Directorate-General for Communications Networks, *Publications Office* 20, 2019, <https://data.europa.eu/doi/10.2759/34672020>.
- Guidelines on artificial intelligence and data protection, T-PD(2019)01, Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data (Convention 108), Directorate General of Human Rights and Rule of Law, 2019, <https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8>.
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 - Artificial Intelligence Act, *Official Journal of the European Union* L series, 12. 7. 2024, (1), (4), (48).
- Европска конвенција за заштиту људских права и основних слобода, Службени лист СЦГ – Међународни ујовори, бр. 9/03, 5/05 и 7/05 – испр. и Службени гласник РС – Међународни ујовори, бр. 12/10.

ПРАКСА ЕВРОПСКОГ СУДА ЗА ЉУДСКА ПРАВА

- Clarke v. the UK*, бр. представке 23695/02, одлука ЕСЉП од 25. августа 2005.
- Daktaras v Lithuania*, бр. представке 42095/98, пресуда ЕСЉП од 10. октобра 2000.
- Dardanskis and Others v. Lithuania*, бр. представке 74452/13, одлука ЕСЉП од 18. јуна 2019.
- Incal v. Turkey*, бр. представке 22678/93, пресуда ЕСЉП од 9. јуна 1998.
- Kyprianou v. Cyprus*, бр. представке 73797/01, пресуда ЕСЉП од 15. децембра 2005.
- Henryk Urban and Ryszard Urban v. Poland*, бр. представке 23614/08, пресуда ЕСЉП од 28. фебруара 2011.
- Luka v. Romania*, бр. представке 34197/02, пресуда ЕСЉП од 21. јула 2009.
- Miroshnik v. Ukraine*, бр. представке 75804/01, пресуда ЕСЉП од 27. новембра 2008.
- Moiseyev v. Russia*, бр. представке 62936/00, пресуда ЕСЉП од 9. октобра 2008.
- Morice v. France*, бр. представке 29369/10, пресуда ЕСЉП од 23. априла 2015.
- Morreira Ferreira v. Portugal* (по. 2), бр. представке 19867/12, пресуда ВВ ЕСЉП од 11. јула 2017.
- Oleksandr Volkov v. Ukraine*, бр. представке 21722/11, пресуда ЕСЉП од 27. маја 2013.
- Salduz v. Turkey*, бр. представке 36391/02, ст. 50, пресуда ВВ ЕСЉП од 26. новембра 2008.
- Sigurður Einarsson and Others v. Iceland*, бр. представке 39757/15, пресуда ЕСЉП од 4. јуна 2019.
- Vera Fernández-Huidobro v. Spain*, бр. представке 74181/01, пресуда ЕСЉП од 6. јануара 2010.

ИЗВОРИ СА ИНТЕРНЕТА

- https://www.cimplifi.com/resources/transparency-explainability-and-interpretability-of-ai/?utm_source=chatgpt.com.
- <https://www.merriam-webster.com/dictionary/black%20box>.
- <https://www.sciencedirect.com/topics/chemical-engineering/deep-learning>.
- <https://www.sciencedirect.com/topics/computer-science/artificial-intelligence>.
- <https://www.sciencedirect.com/topics/engineering/learning-systems>.
- <https://www.merriam-webster.com/dictionary/machine%20learning>.

Jelena S. Milenković, PhD*

THE USE OF ARTIFICIAL INTELLIGENCE FOR RECIDIVISM RISK ASSESSMENT IN CRIMINAL PROCEEDINGS AND THE RIGHT TO A FAIR TRIAL UNDER ARTICLE 6(1) OF THE EUROPEAN CONVENTION ON HUMAN RIGHTS

Summary

In contemporary criminal justice systems, complex artificial intelligence (AI) models are increasingly utilized to assess the risk of recidivism, thereby rendering crime prevention instruments algorithmic in nature.

The paper examines the concept of recidivism as a form of crime prevention and the instruments employed for assessing recidivism risk within criminal proceedings, with particular emphasis on tools that leverage AI and machine learning to assist judges in decision-making processes. Such AI models are already in use in the United States.

Taking into account the issues that have emerged in practice during the application of these models in the U.S., as well as the general challenges recognized in the use of AI, this paper analyses these problems from the perspective of upholding the guarantees of the right to a fair trial under Article 6, paragraph 1 of the European Convention on Human Rights. The objective is to determine whether such models, in their current form, can be implemented in the member states of the Council of Europe in accordance with the accepted standards of fairness in criminal proceedings.

Keywords: artificial intelligence, recidivism risk assessment, criminal proceedings, right to a fair trial, European Convention on Human Rights.

* State Attorney's Office of the Republic of Serbia, The Agency Sector before the European Court of Human Rights, jelenasmil@yahoo.com.