Pregledni naučni rad
UDK:343.21:004.8
DOI:10.7251/ZND2601029N

RAZOTKRIVANJE NITI UMJETNE INTELIGENCIJE U KRIVIČNOM PRAVOSUĐU

***Filip Novaković, mag. iur.*[5]**
Pravni fakultet Univerziteta u Zenici
Bosansko-hercegovačko-američka akademija umjetnosti i nauka

***Nataša Ranković*[6]**
Pravni fakultet Univerziteta u Beogradu

Sažetak: *U proteklim godinama, susret umjetne inteligencije (AI) i krivičnog pravosuđa potakao je paradigmatski pomak u upravljanju pravdom. Ovaj rad istražuje mnogobrojne implikacije integracije umjetne inteligencije u pravosudni pejzaž. Rad kritički ispituje primjenu AI u krivičnim sudovima, fokusirajući se na njegovu ulogu u procesima donošenja odluka, pristranosti i integriteta. Iskorištavanjem algoritama mehaničkog učenja, ovi sistemi analiziraju obimne skupove podataka kako bi predvidjeli buduće kriminalno ponašanje, postavljajući relevantna pitanja o pouzdanosti i pravičnosti takvih predviđanja. Etičke dimenzije upotrebe AI u krivičnom pravosuđu čine središnju temu, razmatrajući pitanja transparentnosti, odgovornosti i pristranosti inherentne u algoritamskom donošenju odluka. Ovaj tekst analizira izazove i prilike koje pruža „mehanička pravda", naglašavajući potrebu za nijansiranim razumijevanjem uticaja AI na temeljna načela krivičnog postupka i ljudska prava i osnovne slobode. Nadalje, istraživanje analizira potencijal za pogoršanje postojećih dispariteta unutar krivičnog pravosudnog sistema. Društveno-ekonomske, rasne, spolne i rodne pristranosti ukorijenjene u istorijskim skupovima podataka mogu se održavati i čak pojačavati kroz primjenu AI alata, što zahtijeva sveobuhvatne mjere opreza radi ublažavanja takvih rizika. Dok stojimo na raskrižju tehnološkog napretka i pravne tradicije, ovaj rad služi kao ključni resurs za pravne teoretičare, praktičare i kreatora politika. Zagovara promišljen i mjerodavan pristup integraciji AI u krivično pravosuđe, osiguravajući da tehnološki napretci usklađuju s načelima pravičnosti, pravde i zaštite individualnih prava. Rasprava*

[5] filip.novakovic.acc@gmail.com; ORDIC iD: https://orcid.org/0000-0001-9979-5824.
[6] natasarankovic24@gmail.com; ORCID iD: https://orcid.org/0009-0000-9420-401X.

predstavljena ovdje ključna je za oblikovanje kontura budućnosti u kojoj se UI i pravda stapaju, zahtijevajući budno nadziranje i etičko razmatranje u traganju za pravednijim pravnim sistemom.

Ključne riječi: *umjetna inteligencija, krivično pravosuđe, pravičan postupak, sudsko prosuđivanje, pravičnost.*

Uvod

Susret umjetne inteligencije (engl. *Artificial Intelligence – AI*) i pravosuđa označava ključni trenutak u evoluciji pravnih sistema širom svijeta. Dok stojimo na rubu ere u kojoj mašine sve više učestvuju u procesima donošenja odluka unutar pravosudnog sistema, postaje imperativno da pravni stručnjaci provode rigorozno ispitivanje implikacija, izazova i etičkih konsideracija koje prate integraciju AI tehnologija. Ovaj članak nastoji istražiti kompleksan pejzaž „mehaničkog“ pravosuđa, zaronivši u višeslojne niti koje prolaze kroz tkaninu naših krivičnih sudova.

Dolazak AI tehnologija u pravnu sferu „obećao“ je neviđene napretke u efikasnosti, preciznosti i objektivnosti. Algoritmi mašinskog / mehaničkog učenja, vođeni ogromnim skupovima podataka i sofisticiranim računarskim modelima, mogu se koristiti kako bi pomogli pravnim stručnjacima u različitim aspektima krivičnog pravosuđa, od prediktivnog izvršavanja zakona do preporuka za izricanje kazni. Međutim, ove primjene ne dolaze bez svojih nijansi i zabrinutosti, što nas potiče da proučimo delikatnu ravnotežu između tehnološke inovacije i očuvanja temeljnih principa prava i krivičnog pravosuđa.

Naime, umjetna inteligencija, iako se rapidno razvija i nalazi primjenu u različitim sektorima, suočava se sa značajnim ograničenjima u kontekstu primjene u pravnom sistemu, posebno kada je riječ o kreiranju novih pravnih normi, praksi i rešavanju pravnih pitanja koja zahtijevaju inovaciju. Uvođenje umjetne inteligencije u pravni okvir podrazumijeva prepoznavanje postojećih normi, pravila i sudskih odluka, ali se ona oslanja na već postojeće informacije, što znači da je njen kapacitet za generisanje novih, kreativanih pravnih rešenja značajno ograničen.

U središtu paradigme mehaničkog pravosuđa leži pitanje pravičnosti. Kako povjeravamo zadatke algoritmima koji su tradicionalno obavljani od strane ljudi, moramo se suočiti s inherentnim pristrasnostima prisutnim i u podacima korištenim za obuku ovih sistema i u samim algoritmima. Koncept „algoritamske pristrasnosti“ postavlja temeljna pitanja o potencijalnom održavanju ili

pogoršanju postojećih dispariteta unutar krivičnog pravosuđa. Kako osigurati da algoritmi, često smatrani neutralnim entitetima, ne nametnu nenamjerno i pojačaju društvene predrasude?

Osim toga, upotreba AI u krivičnim sudovima zahtijeva ponovnu procjenu tradicionalnih pojmova odgovornosti i transparentnosti. Za razliku od ljudskih donosilaca odluka, algoritmi djeluju kao crne kutije, čineći njihove procese donošenja odluka neprozirnima i teškim za tumačenje. Ovaj nedostatak transparentnosti izaziva zabrinutost zbog prava na pravično suđenje, jer optuženi imaju pravo razumjeti osnovu odluka koje utječu na njihove živote. Kako možemo pomiriti potrebu za transparentnošću sa svojstvenom prirodom mnogih AI algoritama, i koje mehanizme možemo implementirati kako bismo osigurali odgovornost kada su ulozi tako visoki kao individualna sloboda?

Dok putujemo kroz pejzaž „mehaničkog“ pravosuđa, etičke konsideracije veličanstveno se nadvijaju nad horizontom. Uključivanje AI tehnologija u krivične postupke zahtijeva pažljivo ispitivanje etičkih implikacija koje okružuju pitanja poput privatnosti, saglasnosti i potencijalne prekomjerne zavisnosti o tehnologiji na štetu ljudskog prosuđivanja. Postizanje ravnoteže između koristi efikasnosti i zaštite individualnih prava ostaje značajan izazov.

Ovaj rad ima za cilj odmotati intrigatne niti koje povezuju AI i krivično pravosuđe, proučavajući ne samo potencijalne zamke i etičke dileme, već i transformacijske mogućnosti koje tehnologija pruža. Kritičkom procjenom trenutnog stanja u pravosuđu i razvoja AI tehnologija, možemo zacrtati put prema budućnosti u kojoj integracija AI-a usklađuje s temeljnim načelima pravde, pravičnosti i zaštite individualnih sloboda. Kako se pravni stručnjaci, praktičari i inženjeri susreću u ovom interdisciplinarnom dijalogu, započinjemo putovanje kako bismo raspetljali složenosti mehaničkog pravosuđa i oblikovali pravni pejzaž koji je istovremeno tehnološki napredan i suštinski pravedan.

Presjek razvoja AI tehnologija i trenutno stanje

Koncept umjetne inteligencije datira još iz drevnih civilizacija, gdje su mitološke priče često prikazivale umjetna bića obdarena ljudskim osobinama. Međutim, formalni početak AI kao naučne discipline može se pratiti do sredine 20. stoljeća. Temeljni period AI svjedočio je značajnim doprinosima ličnosti poput Alana Turinga, koji je predložio Turingov test 1950. godine, i Johna McCarthyja, koji je skovao izraz „umjetna inteligencija“ 1956. godine. Tokom ovog razdoblja, fokus je bio na simboličkoj AI, koristeći logiku i pravila temeljena na sistemima kako bi emulirali ljudsku spoznaju. Unatoč početnom

optimizmu, ovo polje se suočilo s poteškoćama tokom tzv. Zime za AI – razdoblja obilježenog opadajućim interesom i financiranjem (1970-ih i 1980-ih godina) (Russell & Norvig, 2021, pp. 2-3, 16-22; Newquist, 1994, pp. 189-201). Nerealna očekivanja i ograničene računske mogućnosti dovele su do privremenog zastoja u istraživanju AI. Ponovni interes za AI potaknut je napretkom u mehaničkom učenju, posebno neuronskim mrežama tokom 1990-ih i 2000-ih. Usvajanje statističkih pristupa i dostupnost većih skupova podataka i računarske snage omogućili su proboje u područjima poput obrade prirodnog jezika i računalnog vida.

Početak 2010-ih označio je transformacijsko razdoblje u AI obilježeno uzdizanjem arhitektura dubokog učenja. Neuronske mreže, posebno duboke neuronske mreže s više slojeva, pokazale su izvanrednu sposobnost automatskog učenja složenih uzoraka i reprezentacija iz podataka. Ovaj proboj, zajedno s proširenjem *big data* i napretkom u paralelnom računanju, značajno je ojačao sposobnosti AI sistema. Tokom ovog razdoblja, modeli dubokog učenja poput konvolucijskih neuronskih mreža (*convolutional neural networks – CNN*) i rekurentnih neuronskih mreža (*recurrent neural networks – RNN*) revolucionirali su računalni vid, obradu prirodnog jezika i prepoznavanje govora. Klasifikacija slika, detekcija objekata i prijevod jezika postigli su neviđenu preciznost, često nadmašujući ljudske performanse u određenim domenama.

Dolazak generativnih modela, kao što je uvođenje Generativnih suparničkih mreža (*Generative Adversarial Networks – GAN*) koje su predstavili Ian Goodfellow i njegovi kolege 2014. godine, omogućio je sintezu realnih podataka, otvarajući put za proboje u generaciji slika i transferu stila. Ovaj razvoj imao je duboke implikacije za kreativne primjene, poput generiranja umjetnosti i kreiranja sadržaja. Osim toga, učenje jačanja, paradigma gdje agenti uče interagirajući sa svojim okruženjem i primajući povratnu informaciju u obliku nagrada, steklo je ugled. Značajna postignuća uključuju trijumf AlphaGo, *deep reinforcement learning* sistema razvijenog od strane DeepMinda, nad ljudskim prvacima u igri Go, prikazujući sposobnost AI da ovlada složenim strateškim igrama.

Uspon računalnih platformi u oblaku i *open-source* okvira poput TensorFlowa i PyTorcha demokratizirao je pristup sofisticiranim AI alatima, potičući živahnu zajednicu istraživača, programera i praktičara. Ovaj saradnički ekosistem doprinio je brzom širenju znanja i ubrzanju inovacija. Unatoč ovim napretcima, pojavili su se izazovi, uključujući zabrinutosti oko interpretacije i objašnjivosti modela dubokog učenja, kao i pitanja vezana uz pristranost u AI algoritmima. Istraživači i praktičari počeli su istraživati načine za ublažavanje

pristranosti, promicanje transparentnosti i osiguranje odgovornog implementiranja AI tehnologija.

Procjena integracije umjetne inteligencije u kazneno pravosuđe

Primjena umjetne inteligencije u krivičnom pravosuđu može biti, uslovno rečeno, neograničena (od analize digitalnih dokaza, slučajeva vještačenja, procjene rizika, pa sve do određivanja jamstva ili pritvora i odmjeravanja kazne). Upravo iz toga razloga predmet naše analize će biti ograničen isključivo na primjenu umjetne inteligencije u postupku pred krivičnim sudom, odnosno tokom suđenja.

Suđenje (ili prema domaćoj terminologiji – glavni pretres) središnja je faza glavnog postupka, čiji je cilj svestrano razrješenje krivične stvari, odnosno utvrđenje postojanja krivičnog djela, njegovog počinitelja, utvrđivanje krivične odgovornosti na strani počinitelja i njegovo adekvatno kažnjavanje. Ova faza krivičnog postupka je kritična jer obiluje mnoštvom procesnih garanacija po stranke u postupku (naročito za okrivljenog), a gdje sudija igra ključnu ulogu u osiguravanju poštenog i pravednog postupka.

Transparentnost i odgovornost

Primjena umjetne inteligencije u krivičnim postupcima postavlja značajne izazove u vezi s transparentnošću i odgovornosti. Priroda AI algoritama, posebno modela mehaničkog učenja, često uključuje kompleksne računske operacije i složene obrasce koji nisu lako dešifrirani od strane neeksperata. Nedostatak transparentnosti postavlja fundamentalno etičko pitanje u vezi s objašnjivosti odluka generisanih od strane AI, što je ključno u pravnom kontekstu gdje je odgovornost od suštinske važnosti.

Jedna od ključnih zabrinutosti je inherentna neprozirnost mnogih AI modela, često nazivana problem „crne kutije“ (engl. *Black Box Problem*). Kada AI sistemi donose odluke, bilo da se radi o predviđanju recidivizma, ocjenjivanju dokaza ili preporučivanju kazni, od suštinske je važnosti da pravni stručnjaci i zainteresovane strane razumiju kako su ove odluke donesene. Nedostatak transparentnosti izaziva sumnje u pravičnost procesa i dovodi u pitanje samu osnovu prava okrivljenika da razumije dokaze koji su mu predočeni (Inspirit AI, n.d.). Temeljna načela krivičnog postupka zahtijevaju da odluke koje utiču na slobodu pojedinca budu otvorene za ispitivanje i osporavanje. U odsustvu jasnog razumijevanja kako AI dolazi do svojih zaključaka, pravni stručnjaci mogu se

pronaći pred izazovom u efikasnom unakrsnom ispitivanju, osporavanju ili žalbi na dokaze generisane od strane AI. Nedostatak objašnjivosti može podrivati antagonistički sistem, koji se oslanja na sposobnost obje strane da efikasno prezentuju i osporavaju dokaze na glavnom pretrenu.

Algoritamska odgovornost je tijesno povezana s transparentnošću sistema umjetne inteligencije, posebno u kontekstu njihove integracije u pravosudni sistem. Ovisnost o istorijskim podacima za obuku AI sistema unosi značajnu etičku zabrinutost, jer ovi sistemi nehotice mogu perpetuirati postojeće pristrasnosti u podacima. Prepoznavanje i ispravljanje pristrasnih rezultata od izuzetne je važnosti, ističući etičku odgovornost koja leži na pravničkoj profesiji. Da bi efikasno upravljali ovim područjem, pravni stručnjaci moraju biti opremljeni snažnim alatima koji omogućavaju identifikaciju i ispravku pristrasnih algoritama. To zahtijeva stalnu posvećenost praćenju, procjeni i prilagodbi AI sistema kako bi se smanjio potencijal za diskriminatorske uticaje na raznorazne demografske grupe.

Osiguranje odgovornosti za greške ili pristrasnosti u AI sistemima postaje posebno izazovno u odsustvu transparentnosti. Pravni okviri igraju ključnu ulogu u rješavanju ovog izazova propisivanjem rigoroznih testiranja, revizija i izvještavanja o AI algoritmima. Ove mjere su ključne kako bi se osigurala usklađenost s unaprijed definisanim standardima pravičnosti i tačnosti. Stoga, proaktivno prepoznavanje i ispravljanje pristranosti moraju biti integrisani u razvojne i implementacijske procese AI sistema kako bi se očuvali etički standardi, posebno u kontekstu krivičnih postupaka.

Integracija AI u pravosudni sistem donosi obećanje poboljšane efikasnosti, tačnosti i pravičnosti. Ipak, rastuća važnost AI tehnologija izaziva legitimne zabrinutosti u vezi s pristrasnošću i etičkim razmatranjima. AI sistemi nisu otporni na pristrasnost, jer su njihovi procesi učenja ukorijenjeni u istorijskim podacima koji mogu sadržavati društvene predrasude i sistematske nejednakosti. Korištenje pristrasnih podataka u obuci AI algoritama predstavlja stvarni rizik od održavanja i pojačavanja diskriminatorskih praksi unutar pravosudnog sistema. Kritička evaluacija trening skupova podataka neophodna je kako bi se efikasno prepoznale i ublažile pristrasnosti, čime se sprječava da AI sistemi ojačaju postojeće disparitete. Redovne revizije, oslanjanje na raznolike i reprezentativne skupove podataka te kontinuirano praćenje izlaze kao neophodne prakse za proaktivno rješavanje pristrasnosti (Idder & Coulaux, 2021).

Sama priroda AI algoritama koja nosi neprozirnost predstavlja izazove za transparentnost i odgovornost. Nedostatak transparentnosti podriva povjerenje javnosti i izaziva valjane zabrinutosti u vezi sa pravičnošću odluka koje donosi ili

inicira AI. Kako bi se ublažile ove zabrinutosti, AI sistemi trebaju biti dizajnirani s obzirom na objašnjivost, omogućavajući korisnicima razumijevanje procesa donošenja odluka. Usvajanje metodologija objašnjive AI olakšava provjeru i odgovornost, pružajući uvid u to kako se dolazi do zaključaka. Dodatne mjere, uključujući objavu smjernica, otkrivanje izvora trening podataka i pristupačnost algoritamskih detalja, doprinose rješavanju pristrasnosti i osiguravanju pravičnosti.

Iako AI može služiti kao vrijedan alat u procesima donošenja odluka, konačna odgovornost leži na ljudima. Održavanje ljudske nadzora u pravosudnom sistemu je imperativ, predstavljajući AI kao sredstvo za nadopunu ljudskog prosuđivanja umjesto zamjene. Bez obzira na autonomnu prirodu operacija AI, prepoznavanje njenog ljudskog porijekla naglašava mogućnost grešaka. Svi skupovi podataka koji se koriste u AI algoritmima sadrže inherentne ljudske pristrasnosti, ističući potrebu za savjesnim nadzorom.

Da bi se riješili izazovi transparentnosti i odgovornosti, pravni sistemi trebaju uspostaviti jasne standarde i regulacije koje upravljaju upotrebom AI u krivičnim postupcima. Eticki smjernice moraju biti razvijene kako bi se definisao dopušteni opseg i ograničenja primjene AI, naglašavajući važnost transparentnosti u algoritamskom donošenju odluka. Regulatorna tijela trebaju odigrati proaktivnu ulogu u nadgledanju implementacije ovih smjernica, osiguravajući da se AI sistemi pridržavaju etičkih standarda i ne kompromitiraju integritet krivičnih suđenja.

Pristrasnost i pravičnosti

Prisustvo pristrasnosti unutar algoritama umjetne inteligencije predstavlja ključnu etičku zabrinutost, posebno kada se primjenjuje u kontekstu krivičnog pravosuđa. Predrasude se mogu manifestovati na različite načine, proizlazeći iz istorijskih neravnoteža u praksama sprovođenja zakona, uzorcima hapšenja ili lišenja slobode, kao i ishodima izricanja kazni. Razumijevanje i umanjivanje pristrasnosti u AI sistemima ključno je za očuvanje tmeljnog principa pravičnosti u krivičnom pravosuđu.

AI modeli uče iz istorijskih podataka, i ako su ti podaci pristrasni, rezultirajući algoritmi mogu održavati i čak pogoršati postojeće nejednakosti. Istorijske nepravde i sistematske pristrasnosti prisutne u podacima krivičnog pravosuđa, poput rasno nesrazmjernih stopa hapšenja ili dispariteta u izricanju kazni, mogu nehotice biti ugrađene u AI sisteme. Kao rezultat toga, upotreba

pristrasnih AI u krivičnim suđenjima može ojačati diskriminatorske prakse, nerazmjerno utječući na određene demografske grupe (Gardner, n.d.).

Rješavanje pristrasnosti podataka zahtijeva pažljivu analizu trening skupova podataka kako bi se identificirale i ispravile istorijske neravnoteže. Pravni stručnjaci i razvijatelji AI moraju sarađivati kako bi osigurali da su trening podaci reprezentativni i slobodni od diskriminirajućih obrazaca. Štaviše, kontinuirano praćenje i revizija AI sistema su neophodni kako bi se otkrile i ispravile pristrasnosti koje mogu nastati tokom implementacije.

Osiguranje pravičnosti u algoritamskom donošenju odluka kompleksan je zadatak koji zahtijeva pažnju prema različitim dimenzijama jednakosti. Pravni sistemi naglašavaju važnost jednog tretmana pojedinaca pred zakonom, bez obzira na rasu, spol, rod, nacionalnu, etničku ili vjersku pripadnost ili druge zaštićene karakteristike. Međutim, AI sistemi mogu nehotice uvoditi ili održavati disparitete ako nisu dizajnirani i kalibrirani s pravičnošću na umu. Jedan pristup rješavanju problema pravičnosti je razvoj algoritama koji eksplicitno uzimaju u obzir i umanjuju pristrasnost. Tehnike mašinskog učenja svjesne pravičnosti teže balansiranju prediktivne tačnosti između različitih grupa, čime se smanjuje rizik od diskriminatornih ishoda. Pravni okviri trebaju poticati usvajanje takvih tehnika i propisivati evaluaciju AI sistema za pravičnost prije njihove implementacije u krivičnim suđenjima (Hao, 2019).

Etičke implikacije pristrasne AI proširuju se na osnovne principe postupka i jednake zaštite pred zakonom. Ako AI sistemi sistematski oštećuju određene grupe, pojedinci iz tih grupa mogu se suočiti s povećanim rizicima od pogrešnih presuda, nepravičnih kazni ili nejednakog postupanja unutar krivičnog pravosuđa. Ovo ne samo da podriva principe pravde, već i podriva povjerenje javnosti u pravni sistem. Da bi se osigurao postupak i jednaka zaštita, pravni stručnjaci moraju kritički procijeniti pravičnost AI-generiranih rezultata. To uključuje ispitivanje ne samo rezultata koje proizvode AI sistemi, već i metodologija korištenih u njihovom razvoju. Ustanovljavanje jasnih standarda za pravičnost u algoritamskom donošenju odluka i uključivanje tih standarda u pravne okvire može pomoći u smanjenju rizika od nenamjernih diskriminatorskih posljedica (Custers, 2022, pp. 216-217).

U suštini, suočavanje s pristrasnošću i promicanje pravičnosti u upotrebi umjetne inteligencije u krivičnim suđenjima zahtijeva budan i proaktivan pristup. Pravni stručnjaci moraju sarađivati s inženjerima kako bi pažljivo ispitivali podatke, algoritme i rezultate na moguće pristrasnosti. Etički smjernice i pravni standardi trebaju biti uspostavljeni kako bi se osiguralo da AI sistemi doprinose pravičnom i jednakom krivičnom pravosuđu, slobodnom od održavanja

istorijskih nepravdi i sistematskih pristrasnosti. Prioritetiziranjem pravičnosti, integracija AI u krivična suđenja može se više uskladiti s etičkim imperativima pravne profesije.

Zaštita privatnost i sigurnost podataka

Integracija umjetne inteligencije u krivičnim postupcima uključuje prikupljanje, procesuiranje i analizu ogromnih količina osjetljivih ličnih podataka. Pravni postupci inherentno uključuju osjetljive informacije o pojedincima, uključujući detalje o njihovom privatnom životu, kriminalnoj istoriji i potencijalno, mentalnom zdravlju. Upotreba AI u krivičnim suđenjima zahtijeva pažljivo ispitivanje opsega i prirode prikupljanja podataka. Imperativ je uspostaviti jasne smjernice u vezi sa vrstama informacija koje mogu biti korištene od strane AI sistema i osigurati da su pojedinci adekvatno informisani o potencijalnim upotrebama njihovih podataka.

Poštovanje prava pojedinaca na privatnost zahtijeva dobijanje informisanog pristanka za prikupljanje i korištenje ličnih podataka u kontekstu primjene AI. Pravni okviri moraju propisati transparentne prakse otkrivanja, osiguravajući da pojedinci znaju kako će se njihovi podaci koristiti u krivičnim postupcima. Mehanizmi informisanog pristanka osnažuju pojedince da donose odluke o mjeri u kojoj su voljni dopustiti da njihovi podaci utiču na pravne procese donošenja odluka.

Veličanstvene rezervoare podataka koje mogu (ili koje će) da koriste AI sistemi u krivičnim postupcima predstavljaju privlačan cilj za zlonamjerne aktere koji traže neovlašten pristup. Osiguranje snažnih mjera sigurnosti podataka ne samo da je etički imperativ, već i pravna obaveza. Povrede sigurnosti podataka mogu ugroziti povjerljivost i integritet osjetljivih informacija, što može dovesti do ozbiljnih posljedica za pojedince uključene u krivične postupke.

Pravni stručnjaci, u saradnji s ekspertima za informacionu tehnologiju, moraju implementirati i kontinuirano ažurirati stroge protokole sigurnosti kako bi se zaštitili od neovlaštenog pristupa, hakiranja ili povreda podataka. To uključuje enkripciju, kontrole pristupa, redovne revizije i pridržavanje najboljih praksi industrije za sigurnost podataka. Štaviše, pravni okviri trebaju propisivati stroge kazne za neovlašteni pristup i zloupotrebu podataka koje generira AI, ističući ozbiljnost održavanja integriteta pravosudnog sistema.

Održavanje povjerenja javnosti u pravosudni sistem ovisi o uvjeravanju pojedinaca da se njihovi podaci obrađuju s najvećom pažnjom i poštovanjem privatnosti. Svaka percepcija „opuštenih“ praksi privatnosti podataka ili

sigurnosnih ranjivosti može narušiti povjerenje u pravičnost i integritet procesa. Javna svijest i edukacija o mjerama koje su uvedene kako bi se zaštitila privatnost podataka su ključni sastojci izgradnje povjerenja. Pravni okviri trebaju uključivati odredbe za transparentno izvještavanje o praksama sigurnosti podataka, pokazujući posvećenost etičkim standardima i odgovornosti. Prioritetizacijom privatnosti i sigurnosti podataka, pravni stručnjaci mogu pomoći oblikovati narativ koji osigurava javnost da se integracija AI u krivična suđenja provodi s najvećim poštovanjem individualnih prava i privatnosti.

Robusnost i pouzdanost

Pouzdanost AI-generiranih dokaza i procesa donošenja odluka ključna je etička pretpostavka u integraciji umjetne inteligencije u krivično pravosuđe. Osiguravanje da su AI sistemi robusni, otporni na manipulaciju i dosljedno pouzdani od suštinskog je značaja za očuvanje integriteta krivičnog postupka i zaštitu prava uključenih pojedinaca.

Robusnost AI algoritama odnosi se na njihovu sposobnost da proizvode tačne i dosljedne rezultate u različitim scenarijima. Etičke zabrinutosti javljaju se kada su AI sistemi podložni manipulaciji, bilo putem namjernih adverzalnih napada ili nenamjernog iskorištavanja ranjivosti. Adverzalni napadi uključuju namjerno unošenje suptilnih modifikacija u ulazne podatke s ciljem zavaravanja procesa donošenja odluka AI sistema. Pravni stručnjaci moraju obratiti pažnju na rizik od adverzalnih napada sarađujući s ekspertima iz područja kibernetičke sigurnosti i razvoja AI. Razvoj algoritama koji su otporni na manipulaciju i uključivanje zaštitnih mjera protiv adverzalnih napada ključno je. Osim toga, pravni okviri trebaju propisivati postupke testiranja kako bi se ocijenila robusnost AI sistema prije njihove implementacije u krivična suđenja, osiguravajući njihovu pouzdanost u prisustvu potencijalnih prijetnji.

Pouzdanost AI sistema snažno je utemeljena na kvaliteti i reprezentativnosti podataka na kojima su trenirani. Predrasude i netačnosti prisutne u trening podacima mogu rezultirati nepouzdanim i nepravednim ishodima. Etičke pretpostavke zahtijevaju temeljitu analizu trening skupova podataka kako bi se identificirale i ispravile eventualne manjkavosti, osiguravajući da se AI modeli dobro generaliziraju za raznolike scenarije koji se javljaju u krivičnim postupcima. Strategije ublažavanja pristrasnosti trebaju se primijeniti kako bi se riješile nepravilnosti, a kontinuirano praćenje treba provoditi kako bi se identificirale i ispravile eventualne nove probleme. Transparentno izvještavanje o sastavu trening podataka i poduzetim koracima za

rješavanje pristrasnosti doprinosi općoj pouzdanosti AI sistema u krivičnim postupcima.

Moramo napomenuti da nijedan AI sistem (kao i svaki drugi sistem) nije potpuno slobodan od grešaka i otporan na nedostatke, a razumijevanje i upravljanje stopama grešaka sastavni su dio etičke implementacije AI. Pravni stručnjaci moraju uzeti u obzir potencijalne posljedice AI-generiranih grešaka, posebno u krivičnim suđenjima gdje su ulozi visoki. Odgovornost za greške leži na raskrižju tehnologije i pravne odgovornosti. Etičke smjernice trebale bi uspostaviti prihvatljive stope grešaka za AI sisteme korištene u krivičnim postupcima, uzimajući u obzir potencijalni utjecaj na prava i slobode pojedinaca. Osim toga, pravni okviri moraju definisati mehanizme za utvrđivanje odgovornosti kad dođe do grešaka. Ustanovljavanje jasnih linija odgovornosti osigurava da pravni stručnjaci, razvojni stručnjaci za tehnologiju i ostali dionici snose odgovornost za posljedice odluka generiranih od strane AI.

Nadzor i proces donošenja odluka

Integracija umjetne inteligencije u krivično pravosuđe postavlja duboka etička pitanja o ulozi čovjeka u nadzoru nad radom i pri donošenju odluka. Iako AI sistemi mogu pružiti vrijedne uvide i efikasnost, etička imperativa leži u osiguranju da ljudska prosudba ostane centralna u pravnim procesima, pri čemu AI služi kao alat za pomoć, a ne zamjenu za ljudske odlučitelje. Etički pristup uključivanju AI u krivična suđenja naglašava proširenje ljudskog donošenja odluka umjesto potpune automatizacije pravnih procesa. AI bi trebala biti shvaćena kao alat koji pomaže pravnim stručnjacima u procesiranju velikih skupova podataka, identifikaciji uzoraka i povećanju efikasnosti. Ljudski nadzor osigurava da su nijansirani, kontekstualno zavisni aspekti pravnih postupaka uzeti u obzir, sprječavajući odricanje od autoriteta donošenja odluka u korist algoritama. Pravni okviri trebaju eksplicitno definisati ulogu AI kao proširujućeg alata, razjašnjavajući da konačna odgovornost i autoritet počivaju na ljudima – sudijama, tužiocima, advokatima i drugim pravnim stručnjacima. Ovaj pristup održava integritet pravnog sistema dok istovremeno koristi prednosti AI za pojednostavljivanje određenih procesa.

Etički imperativ ljudskog nadzora usko je povezan s tumačivošću i razumljivošću AI-generiranih odluka. Pravni postupci zahtijevaju transparentnost kako bi se osiguralo da sve strane uključene mogu razumjeti i osporavati dokaze iznesene protiv njih. Ako AI sistemi djeluju kao „crne kutije“ bez jasnih objašnjenja njihovih odluka, pravo na pravično suđenje i efikasnu pravnu

odbranu može biti ugroženo. Ovo dovodi u pitanje pravo na obrazloženu sudsku odluku, koje je sastavni dio prava na pravično (pošteno) suđenje. Kako bi se riješila ova zabrinutost, AI sistemi korišteni u krivičnim postupcima trebali bi biti dizajnirani kako bi proizvodili tumačive rezultate. Pravni stručnjaci moraju imati sposobnost razumijevanja faktora koji utiču na AI-generirane odluke i objašnjavanja istih sudijama, procesnim strankama i drugim zainteresiranim stranama. Jasni standardi za razumljivost AI sistema trebaju biti integrirani u pravne okvire kako bi se osiguralo da je transparentnost neupitni aspekt njihove implementacije.

Pravni sistem, kao kompleksan skup pravnih normi i principa, zahtijeva sposobnost prilagođavanja, tumačenja i razvoja novih pravnih rešenja u odgovoru na društvene promene i nove izazove. Ovaj proces evolucije i prilagođavanja nužno traži duboko razumijevanje ljudskih vrijednosti, društvenih potreba i moralnih dilema, što je umjetnoj inteligenciji, koja se temelji na prethodnim podacima i algoritamskim modelima, izuzetno teško dostići. Ona nije u mogućnosti da samostalno stvori nova pravna načela ili donosi inovativna rešenja koja bi mogla postati osnova nove pravne prakse.

S obzirom na ove okolnosti, uvođenje umjetne inteligencije u pravni sistem mora biti pažljivo promišljeno i sprovedeno sa velikim oprezom. Iako može da podrži pravne stručnjake u analizama i pretrazi prethodnih slučajeva, kao i da automatizuje određene administrativne procese, njeno mjesto u primjeni prava treba biti ograničeno. Uzimajući u obzir da umjetna inteligencija ne može donijeti nove pravne norme ili unaprijediti pravni sistem samostalno, neophodno je da ljudski faktor ostane ključni nosilac odgovornosti za donošenje pravnih odluka, kao i za kreiranje novih pravnih rješenja u skladu sa društvenim i etičkim vrijednostima.

Ljudski nadzor ključan je i za očuvanje etičke i pravne odgovornosti u pravosudnom sistemu. Pravne odluke nose značajne posljedice za živote, slobode i prava pojedinaca. Iako AI može pridonijeti vrijednim uvidima, ljudi su, ipak, bolje opremljeni za razmatranje širih pravnih, etičkih i društvenih implikacija njihovih odluka. Pravni okviri trebaju razgraničiti etičke odgovornosti ljudskih donositelja odluka u korištenju AI. To uključuje obavezu kritičkog procjenjivanja AI-generiranih dokaza, razmatranje njihovih ograničenja i osiguranje da odluke usklađuju s pravnim principima i standardima. Ustanovljavanje jasnog okvira za ljudsku odgovornost umanjuje rizik od nepotrebnog oslanjanja na AI i jača etičke temelje pravnog sistema.

Ljudski nadzor ključan je za zaštitu od pristrasnosti i neočekivanih posljedica koje mogu proizaći iz AI-generiranih odluka. Pravni stručnjaci donose

kontekstualno razumijevanje, pravnu stručnost i svijest o društvenim dinamikama koje AI sistemima može nedostajati. Ljudska prosudba je nužna za procjenu pravičnosti i pravde odluka, posebno kada su suočeni s kompleksnim, novim ili dvosmislenim situacijama. Pravni okviri trebaju naglasiti važnost ljudskog nadzora u identifikaciji i ispravljanju pristrasnosti unutar AI sistema. Redovna obuka pravnih stručnjaka o nijansama primjene AI, potencijalnim pristrasnostima i njihovim utjecajima na pravne ishode, može ih osnažiti da primijene informisanu prosudbu i osiguraju da se korištenje AI usklađuje s etičkim i pravnim standardima.

Prema tome, etičke pretpostavke koje se odnose na ljudski nadzor i donošenje odluka u korištenju umjetne inteligencije u krivičnim suđenjima naglašavaju ključnu ulogu pravnih stručnjaka u očuvanju pravičnosti, transparentnosti i odgovornosti pravosudnog sistema. Propisivanjem AI kao proširujućeg alata pod budnim okom ljudske prosudbe, pravni sistemi mogu koristiti tehnološke napretke dok istovremeno podržavaju temeljne principe pravde, pravne etike i ljudskih prava.

Razmišljanje izvan „Crne kutije“

Integracija umjetne inteligencije u krivične pravosudne sisteme nosi veliko obećanje za efikasnost i točnost. Međutim, etički izazovi koje postavljaju pristrasnosti u AI aplikacijama zahtijevaju sveobuhvatan i kooperativan odgovor. Jedna ključna strategija je osigurati da trening podaci korišteni za razvoj AI modela budu raznoliki i reprezentativni za populaciju. Prepoznajući da pristrasnosti često proizlaze iz istorijskih neravnoteža u podacima, trebalo bi uložiti napore kako bi se eliminirale takve disparitete. Poticanjem inkluzivnosti u trening skupovima podataka, AI sustavi mogu biti trenirani da pruže pravednije rezultate.

Implementacija mehanizama kontinuiranog praćenja i evaluacije ključna je za otkrivanje i ispravljanje pristrasnosti kako se one pojavljuju. Redovite revizije AI sistema omogućuju identifikaciju i rješavanje problema, osiguravajući da se tehnologija razvija s nepokolebljivim opredjeljenjem za pravičnost. Prioritetno je i postavljanje transparentnosti i objašnjivosti u AI sistemima. Modeli bi trebali biti dizajnirani tako da pružaju jasna objašnjenja za svoje odluke, omogućujući pravnim stručnjacima, procesnim strankama i javnosti da razumiju i ispituju razloge iza AI-generiranih rezultata.

Provoditi algoritamske revizije važno je za procjenu uticaja AI sistema na različite demografske skupine. To uključuje evaluaciju ishoda kako bi se

identificirale i ispravile eventualne disparitete ili diskriminatorni obrasci, osiguravajući pravičnost među raznolikim populacijama. Istraživanje i implementacija tehnika *machine learning*-a svjesnih pravičnosti još je jedan ključni korak. Ove metode imaju za cilj garantisati da AI sistemi pružaju pravedne i jednake rezultate, sprječavajući pojačavanje postojećih pristrasnosti i promicanje pravde za sve. Poticanje saradnje između pravnih stručnjaka i eksperata za tehnologiju je imperativ. Pravni stručnjaci trebaju aktivno sudjelovati u razvoju i implementaciji AI sistema, pružajući uvide u pravne nijanse i etičke aspekte. Nadalje, ustanovljavanje jasnih etičkih smjernica i standarda za korištenje AI u krivičnom pravosuđu je ključno. Ove smjernice trebaju naglašavati pravičnost, transparentnost i odgovornost, pružajući okvir za razvoj i primjenu etičkih praksi.

Osiguravanje redovite obuke za pravne stručnjake o mogućnostima i ograničenjima AI sistema također predstavlja značajan segmet u prevalozaženju etučkih i pravnih preprema primjene AI tehnologija u krivičnom pravosuđu. To osnažuje pravne stručnjake da kritički procijene AI-generirane dokaze, identificiraju potencijalne pristrasnosti i donose informirane odluke o dopuštenosti i relevantnosti takvih dokaza. Također, poticanje nezavisnih revizija i pregleda AI sistema korištenih u krivičnim postupcima dodaje dodatni sloj provjere. Vanjski nadzor pomaže izgradnji povjerenja u pravičnost i pouzdanost primjene AI tehnologija, pridonoseći transparentnijoj i odgovornijoj implementaciji tehnologije.

Tretiranje razvoja i implementacije AI sistema kao iterativnog procesa temeljnog je značaja. Učenje iz pogrešaka i kontinuirano unaprjeđivanje algoritama na temelju stvarnih ishoda i povratnih informacija od pravnih stručnjaka i pogođenih zajednica poboljšava prilagodljivost i pravičnost primjene AI aplikacija. Prevladavanje etičkih prepreka povezanih s pristrasnostima u korištenju AI u krivičnim suđenjima zahtijeva sveobuhvatan i suradnički pristup. Integracijom pravnih, tehničkih i društvenih perspektiva moguće je razviti i implementirati AI sisteme koji se usklađuju s principima pravde, pravičnosti i jednake zaštite pred zakonom. Cilj je stvoriti pravosudni sistem koji prihvaća prednosti AI dok osigurava nepristrane i pravedne ishode za sve pojedince uključene.

Zaključak

U labirintskom krajoliku modernog pravosuđa, integracija umjetne inteligencije bacila je duboku i transformirajuću sjenu. Kako se diskutovalo kroz ovaj članak, korištenje umjetne inteligencije u krivičnom pravosuđu donosi određeni spektar prednosti i nedostataka, s etičkim implikacijama koje odzvanjaju u središtu ove tehnološke konvergencije. Dok stojimo na raskrižju inovacija i tradicije, imperativ je razmisliti o putu koji slijedimo i pažljivo razmotriti posljedice predaje dijelova krivičnog postupka dominaciji umjetne inteligencije.

Nesumnjivo, prednosti upotrebe alata AI u krivičnim postupcima su značajne. Dobici u učinkovitosti / djelotvornosti, ubrzanje analize podataka i potencijal za nepristrasno donošenje odluka među su najviše hvaljenim aspektima. Algoritmi mehaničkog učenja, kada su pravilno obučeni i implementirani, mogu obraditi ogromne količine informacija u djeliću vremena koje bi bilo potrebno ljudskom kolegi. Izgledi smanjenja zastoja, poboljšanje prodiktivnih sposobnosti i ublažavanje ljudske pogreške čine se obećavajućima. Osim toga, bezstranost AI sistema obećava donošenje presuda oslobođenih osobne pristranosti, što je uvijek prisutna briga u tradicionalnim pravnim postupcima.

No, zov napretka prati kakofonija etičkih dilema. Neprozirnost procesa donošenja odluka putem AI izaziva zabrinutosti u vezi s odgovornošću i transparentnošću. Postavljaju se pitanja o pravičnom postupanju s pojedincima, posebno u kontekstima gdje se istorijske pristranosti mogu nenamjerno održavati algoritamskim modelima. Delikatna ravnoteža između učinkovitosti i poštovanja prava na pravično suđenje, etičke implikacije automatiziranog donošenja odluka i potencijalno narušavanje temeljnih ljudskih prava nadvijaju se nad budućnošću „algoritamskog“ pravosuđa. Kao čuvari pravde, moramo oprezno kročiti, snalazeći se u zamršenom etičkom labirintu koji se pojavljuje s integracijom AI u krivične postupke.

U konačnoj analizi, imperativ očuvanja ljudskog elementa u krivičnom postupku ostaje ključan. Bogatstvo ljudskog iskustva, kapacitet za empatiju i nijansirano razumijevanje konteksta su nezamjenjivi dijelovi pravosudnog sistema. I dok AI može briljirati u obradi podataka i prepoznavanju uzoraka, nema prirodnu sposobnost razumijevanja kompleksnosti ljudskih emocija, motivacija i društvenih nijansi koje često leže u osnovi krivičnih slučajeva. Sudnica je kovčeg u kojem se pričaju ljudske priče, i pravda nije samo funkcija podataka, već duboko razumijevanje ljudskog stanja. Osim toga, sama priroda pravde zahtijeva odgovornost, a odgovornost zahtijeva subjektivnog, moralnog i

savjesnog agenta. Mašina nema intrinzičan osjećaj odgovornosti koji je ključan za pravedan i pošten pravni sistem. Potencijal za neželjene posljedice, neočekivane pristranosti i narušavanje principa pravde prilagođene pojedincu naglašava potrebu za opreznim pristupom. U suštini, dok plovimo neistraženim teritorijama „mehaničke pravde“, moramo biti svjesni delikatne ravnoteže između tehnološkog napretka i očuvanja temeljnih ljudskih vrijednosti. Prednosti AI u krivičnim postupcima su privlačne, ali etički izazovi i nezamjenjiva vrijednost ljudske prosudbe traže oprez. Prema tome, dok umjetna inteligencija može igrati značajnu podršku u tehničkom aspektu pravnog rada, njeno uvođenje u ključne pravne procese, koji podrazumijevaju donošenje novih normi i razvoj pravne prakse, treba biti sprovodeno s velikom rezervom, uz jasno definisane granice i nadzor od strane stručnjaka i zakonodavaca.

Kao pravni teoretičari, pravni praktičari, ali i čuvari pravde, moramo ostati budni u zaštiti integriteta našeg pravosudnog sistema, osiguravajući da se težište pravde i dalje drži rukama sposobnim razabrati zamršeni tapiseriju ljudskih života.

Bibliografija

Crevier, D. (1993). *AI: The tumultuous search for artificial intelligence*. New York, NY: Basic Books.

Custers, B. H. M. (2022). AI in criminal law: An overview of AI applications in substantive and procedural criminal law. In B. H. M. Custers & E. Fosch Villaronga (Eds.), *Law and artificial intelligence*. Heidelberg: Springer.

Gardner, R. (n.d.). *AI in criminal justice: How it can become biased*. Inspirit AI. https://www.inspiritai.com/blogs/ai-student-blog/ai-in-criminal-justice-how-it-can-become-biased (Accessed December 14, 2023).

Hao, K. (2019, January 21). *AI is sending people to jail—and getting it wrong*. MIT Technology Review. https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/ (Accessed December 14, 2023).

Idder, A., & Coulaux, S. (2021, December 13). *Artificial intelligence in criminal justice: Invasion or revolution?* International Bar Association. https://www.ibanet.org/dec-21-ai-criminal-justice#_ednref1 (Accessed December 14, 2023).

Implications of AI in the criminal justice system. (n.d.). Inspirit AI. https://www.inspiritai.com/blogs/ai-student-blog/implications-of-ai-in-the-criminal-justice-system (Accessed December 14, 2023).

Newquist, H. P. (1994). *The brain makers: Genius, ego, and greed in the quest for machines that think*. Indianapolis, IN: Macmillan.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

UNRAVELING THE THREADS OF ARTIFICIAL INTELLIGENCE IN CRIMINAL JUSTICE

Filip Novaković[7]
Faculty of Law of the Univerzity of Zenica
Bosnian-Herzegovinian American Academy of Arts and Sciences

Nataša Ranković[8]
Faculty of Law of the Univerzity of Beograd

Summary: *In recent years, the intersection of artificial intelligence (AI) and criminal justice has prompted a paradigm shift in the administration of justice. This paper explores the numerous implications of integrating artificial intelligence into the legal landscape. The paper critically examines the application of AI in criminal courts, focusing on its role in decision-making processes, bias, and integrity. By harnessing machine learning algorithms, these systems analyze extensive datasets to predict future criminal behavior, raising pertinent questions about the reliability and fairness of such predictions. The ethical dimensions of using AI in criminal justice are a central theme, considering issues of transparency, accountability, and bias inherent in algorithmic decision-making. This text analyzes the challenges and opportunities presented by „mechanical justice", emphasizing the need for a nuanced understanding of the impact of AI on fundamental principles of criminal procedure and human rights and freedoms. Furthermore, the research examines the potential exacerbation of existing disparities within the criminal justice system. Socio-economic, racial, gender, and sex-based biases rooted in historical datasets can be perpetuated and even amplified through the application of AI tools, necessitating comprehensive precautions to mitigate such risks. As we stand at the crossroads of technological advancement and legal tradition, this paper serves as a crucial resource for legal theorists, practitioners, and policymakers. It advocates for a thoughtful and authoritative approach to integrating AI into criminal justice, ensuring that technological advancements align with the principles of fairness, justice, and the protection of individual rights. The discussion presented here is essential for shaping the contours of a future where AI and justice converge, requiring vigilant oversight and ethical consideration in the pursuit of a more just legal system.*

Keywords: *artificial intelligence, criminal justice, fair procedure, judicial decision-making, justice.*

[7] filip.novakovic.acc@gmail.com; ORDIC iD: https://orcid.org/0000-0001-9979-5824.
[8] [8] natasarankovic24@gmail.com; ORCID iD: https://orcid.org/0009-0000-9420-401X.