

REZOLUCIJE AIDP-a 2024.

XXI. MEĐUNARODNI KONGRES KAZNENOG PRAVA O UMJETNOJ INTELIGENCIJI I KAZNENOM PRAVOSUĐU

Pariz, Francuska, 25.–28. lipnja 2024.

I. SEKCIJA: TRADICIONALNI INSTITUTI KAZNENOG PRAVA I UMJETNA INTELIGENCIJA *

Rezolucija je usvojena na kolokviju „Tradicionalne kategorije kaznenog prava i umjetna inteligencija: kriza ili palingeneza?“, Siracusa, Italija, 14.–17. rujna 2022.

Preambula

Budući:

- da sustavi umjetne inteligencije (UI) s različitim stupnjevima autonomije podržavaju i zamjenjuju mnoge ljudske aktivnosti;
- da bi sustavi UI-ja mogli predstavljati stvarnu korist za društvo u cjelini i za tijela za provođenje zakona, posebno kada je riječ o istragama kaznenih djela;
- da sustavi UI-ja postaju sve autonomniji, a njihovo funkcioniranje može biti nepredvidivo za one koji ih dizajniraju, programiraju, proizvode, distribuiraju i koriste;

Primjećujući:

- da je područje primjene sustava UI-ja vrlo široko, a nezakonite radnje vezane uz njihovu primjenu mogu štetiti raznim interesima, pravnim dobrima i temeljnim pravima;
- da sustavi UI-ja također mogu imati sve veću ulogu u činjenju kažnjivih radnji kao „alati“ za počinjenje kaznenih djela i mogli bi postati čimbenik koji olakšava nastanak novih kaznenih djela;

* Rezoluciju je preveo Nikša Vojvoda, mag. iur., asistent u statusu doktoranda na uspostavljenom istraživačkom projektu Hrvatske zaklade za znanost „Hrvatsko prekršajno pravo u europskom kontekstu – izazovi i perspektive“ (UIP-2020-02-6482) na Katedri za kazneno procesno pravo Pravnog fakulteta Sveučilišta u Zagrebu. Prevoditelj se kao pomoćnim alatom za prevođenje koristio platformama UI-ja *ChatGPT* i *Google Translate*. Prijevod je temeljito provjeren, ispravljen, kontekstualiziran i usklađen sa standardnom pravnim terminologijom.

Obraćajući posebnu pozornost:

- na sve veće potpuno ili djelomično delegiranje odluka na sustave UI-ja u različitim područjima djelovanja, što postavlja pitanje odgovornosti fizičkih ili pravnih osoba za štetu uzrokovanu autonomnim funkcioniranjem takvih sustava;
- na autonomiju sustava UI-ja i mogućnosti o kojoj se raspravlja unutar akademske zajednice da se oni smatraju počiniteljima kaznenih djela;

Imajući u vidu:

- važnost odgovarajućih reakcija koje kazneno pravo treba osigurati za sprječavanje i kažnjavanje kaznenih djela počinjenih od strane, putem ili protiv sustava UI-ja;
- ozbiljnost štete i rizika povezanih s primjenom UI-ja;
- temeljna načela koja moraju biti zajamčena pri utvrđivanju i primjeni kaznenih sankcija (uključujući i kaznene sankcije u širem smislu, koje bi se mogle primijeniti na pravne osobe), kao što su načelo zakonitosti i načelo krivnje, koje su nužan izraz osobne kaznene odgovornosti;
- da treba razmotriti tradicionalne institute kaznenog prava i modele kaznene odgovornosti te ih, ako je potrebno, prilagoditi novonastalim zahtjevima za zaštitom;

Uzimajući u obzir:

- „Etičke smjernice za pouzdani UI“, koje je Stručna skupina visoke razine predstavila Europskoj komisiji 8. travnja 2019., i druge značajne preporuke međunarodnih tijela (na primjer „Studiju izvedivosti o budućem instrumentu Vijeća Europe o umjetnoj inteligenciji i kaznenom pravu“ Europskog odbora za probleme kriminaliteta Vijeća Europe, 4. rujna 2020.);
- prijedlog europske Uredbe o umjetnoj inteligenciji (tzv. Akt o UI-ju); rad Odbora za umjetnu inteligenciju Vijeća Europe; aktivnosti Ujedinjenih naroda na području umjetne inteligencije;
- preporuke XIV. Međunarodnog kongresa (Beč, 1989.) o pravnim i praktičnim problemima koji proizlaze iz razlike između kaznenog i upravno-kaznenog prava; preporuke s XVIII. Međunarodnog kongresa (Istanbul, 2009.) o inkriminaciji pripremanja i sudjelovanja u kaznenom djelu; i one s XIX. Međunarodnog kongresa (Rio de Janeiro, 2014.) o informacijskom društvu i kaznenom pravu;

Sudionici međunarodnog kolokvija I. sekcije XXI. Međunarodnog kongresa kaznenog prava (Kazneno pravo: opći dio): „Tradicionalne kategorije kaznenog prava i UI: kriza ili palingeneza?“ usvojili su sljedeće:

REZOLUCIJE

1. O pojmu umjetne inteligencije i dodjeli pravne osobnosti sustavima umjetne inteligencije s različitim stupnjevima autonomije

1. Pojam UI-ja obuhvaća višestruke algoritamske i robotske sustave koji su u interakciji s okolinom, koji su razvijeni s pomoću nekoliko tehnika (kao što je strojno učenje) i koji služe za postizanje ciljeva koje su definirali ljudi. Stoga za svrhe kaznenog prava nije poželjno ponuditi opću definiciju UI-ja.

2. Međutim, budući da sustavi UI-ja mogu biti štetni u raznim područjima (npr. autonomna vozila, robotski sustavi u medicini, trgovinski sustavi UI-ja ili upravljanje logistikom), kaznenopravna zaštita pravnih dobara i temeljnih prava trebala bi uzeti u obzir specifične značajke različitih sustava UI-ja s različitim stupnjevima autonomije, kao i pravne definicije koje pružaju izvori izvan kaznenog prava u svakom pojedinom području.
3. Prema trenutačnom stanju nema normativnog temelja ni dosljednosti u pogledu funkcije kaznenopravnog sankcioniranja u priznavanju pravne osobnosti sustava UI-ja s različitim stupnjevima autonomije.
4. S jedne strane postoji ontološka razlika u odnosu na ljude kao počinitelje. Sustavima UI-ja nedostaje svijest u odabiru i procjeni mogućih rješenja za problem ili dilemu, uzimanje u obzir konteksta društvenih i etičkih odnosa i prilika, potrebna fleksibilnost i sposobnost prilagodbe čak i izvanrednim ili neočekivanim situacijama i uvjetima.
5. S druge strane, punitivne sankcije primijenjene na takve tehnološke sustave i subjekte ne bi odgovarale svrsi i funkciji kažnjavanja jer bi učinak prijetnje kaznom i njezinom primjenom iščeznuo zbog nedostatka svijesti o vlastitom postojanju u prošlosti, sadašnjosti i budućnosti, a prije svega zbog odsutnosti svjesnog samoodređenja, tako da, sve i da se isključi retributivna funkcija, ne bi bile ostvarive čak ni funkcije specijalne i generalne prevencije.

2. O potrebi izvankaznenog uređenja, standarda i obveza

6. Kako bi se spriječile i smanjile štete povezane s UI-jem prije ili barem usporredno s reformom kaznenog prava, bilo bi potrebno da međunarodni, regionalni i nacionalni zakonodavci i nadležna tijela u potpunosti reguliraju nekoliko sektora u kojima se implementiraju sustavi UI-ja (poput prije spomenutih sektora autonomnih vozila, zdravstvenih i kirurških robota, autonomnog oružja itd.). Trebalo bi regulirati tehničke standarde, strukturne karakteristike i uvjete rada sustava UI-ja i njihovih komponenti.
7. Takvi propisi, koji moraju djelovati od faze dizajna, proizvodnje, distribucije i prodaje do stvarne upotrebe sustava UI-ja, također bi trebali predvidjeti konkretne uvjete za prilagodbu za slučaj upozorenja (crvenih zastavica; *red flags*) ili signala kao preduvjeta za rješavanje problema štete povezane s UI-jem putem punitivnog prava.
8. Ti propisi mogu predvidjeti postupke zabranjivanja, poput onih koji već postoje u područjima složenih rizika (npr. zdravlje i sigurnost na radnom mjestu i zaštita okoliša), čije kršenje ili nepoštovanje može biti sankcionirano prema načelu *ultima ratio*.

3. O potrebi kaznenopravne zaštite pravnih dobara

9. Treba prepoznati ključnu važnost razumne i razmjerne intervencije kaznenog prava u širem smislu u sprječavanju i kažnjavanju šteta i opasnosti koje bi sustavi UI-ja mogli prouzročiti interesima, pravnim dobrima i temeljnim pravima ako bi u jednakim činjeničnim okolnostima da su ih uzrokovale fizičke ili pravne osobe prema tradicionalnim institutima kaznenog prava predstavljale kaznena djela. Stoga djelo ne može proći nekažnjeno samo zato što su ga počinili takvi sustavi, ili su djela počinjena pomoću njih ili protiv njih.

10. Potrebno je identificirati i definirati specifične modele za pripisivanje odgovornosti osobama (i fizičkim i pravnim) koje „stoje iza“ sustava UI-ja (tj. dionici- ma u različitim fazama životnog ciklusa sustava UI-ja, kao što su projektanti, dobavljači, uvoznici, distributeri, korisnici itd.), počevši od vlasnika i onih koji odlučuju o njihovoj konkretnoj uporabi, na temelju svojeg interesa i svoje kori- sti, i koji stoga moraju biti pravno odgovorni, uključujući i „punitivnu“ – a ne samo kaznenu – perspektivu.
11. Odgovornost osoba opisanih u prethodnoj točki ne isključuje odgovornost dru- gih osoba (bilo fizičkih bilo pravnih) koje pridonose uzročnom lancu štete, od dizajnera, programera, proizvođača, prodavača i distributera do krajnjih kori- snika samih sustava.
12. Posebno treba praviti razliku između:
 - a. sustava UI-ja koji se koriste u nezakonitim aktivnostima. U ovom području nepoželjna ponašanja stvaraju manje izazova u pogledu pripisivanja kaznene odgovornosti s obzirom na to da se sustavi UI-ja konceptualno ne razlikuju od ostalih instrumenata i sredstava počinjenja kaznenog djela. Potrebno je, međutim, razmotriti dva pitanja:
 - a.1. U slučaju odstupajućih rezultata funkcioniranja sustava od namjerava- ne nezakonite radnje moraju se primijeniti klasični instituti *aberratio ictus* i *aberratio delicti*. Puka materijalna različitost oštećenog pred- meta ne smije predstavljati razlog za isključenje odgovornosti ako nje- gova svojstva nisu relevantna za ostvarenje elemenata kaznenog djela (npr. ubijanje jedne osobe umjesto druge nije relevantno za ostvarenje kaznenog djela ubojstva kada je počinitelj to namjeravao). Umjesto toga, poželjnije je kaznenu odgovornost za kazneno djelo različito od onoga koje se namjeravalo počiniti temeljiti na mogućnosti konkret- ne predvidljivosti takva različitog razvoja radnji, koje provodi sustav UI-ja primjenom načela odgovornosti temeljene na nehaju (kao što je navedeno u poglavlju 4. u nastavku).
 - a.2. Budući da se sustavi UI-ja mogu koristiti za posebno štetna ili opasna ponašanja, oni mogu pojačati i otežati prouzročenu štetu (kao što se događa s IKT-om).¹ Posljedice doista mogu biti vrlo udaljene od radnji koje su ih uzrokovale, što otežava intervenciju *post factum* radi spre- čavanja ili barem okončavanja ili smanjenja štetnih događaja. Stoga treba razmotriti inkriminaciju protuzakonitog dizajniranja, programi- ranja, proizvodnje, distribucije, prodaje i kupnje „štetnih“ algoritama, softvera i sustava UI-ja kao samostalnih pripremnih djela. Ova kazne- na politika trebala bi biti ograničena na sustave UI-ja koji predstavljaju visok rizik za određena važna pravna dobra (kao što su život, tijelo ili sloboda drugih ljudskih bića), i to samo na slučajeve jasne, stvarne, sadašnje opasnosti (pod uvjetima potrebnima za inkriminiranje pri- premnih akata, vidi rezoluciju I. sekcije XVIII. kongresa AIDP-a u Istanbulu, 2009.).

¹ IKT – informacijsko-komunikacijska tehnologija.

- b. Sustavi UI-ja koji se koriste za zakonite aktivnosti. Ovo otvara najdelikatnija pitanja u području „dopuštenog rizika“, čije bi se granice trebale urediti željenom regulacijom posebnih sigurnosnih obveza i pravila opreza koja će se primjenjivati na aktivnosti dizajna, razvoja, proizvodnje, distribucije i prodaje, odnosno korištenja sustava UI-ja. Prilagodba modela kaznene odgovornosti u ovom području mora se baviti napetošću koja se može stvoriti između oblika odgovornosti za nesavjesno ponašanje i tehničkih značajki sustava UI-ja, i to: (1) njihove autonomije; (2) konkretne nepredvidivosti njihovih odluka i funkcioniranja; (3) netransparentnosti njihovih regulatornih mehanizama; (4) složenosti procesa njihova programiranja, razvoja, proizvodnje, ažuriranja i održavanja.

4. O prilagodbi modela pripisivanja odgovornosti značajkama sustava UI-ja, konkretno stupnju njihove autonomije

13. Prije svega, u skladu s priznatom postupnom automatizacijom i autonomijom aplikacija UI-ja u nekoliko područja, potrebno je razlikovati razine odlučivanja i operativne autonomije tih sustava, od onih u kojima „automatsko“ funkcioniranje omogućuje ljudskom subjektu da ima značajnu kontrolu nad sustavom do onih koji su zaista „autonomni“, u kojima ljudska intervencija može biti samo vremenski i prostorno udaljena od funkcioniranja sustava UI-ja, koji „odlučuje“ na temelju prikupljenih informacija i algoritama koji se prilagođavaju njegovu iskustvu, tako da postoji strukturna margina nepredvidivosti konkretnih ishoda.
14. U odnosu na različite vrste sustava UI-ja od ključne je važnosti definiranje posebnih pravila i standarda funkcioniranja, kako je predviđeno u prijedlogu europske Uredbe o UI-ju (vidi 2. poglavlje).
15. Najhitnija potreba za prilagodbom tradicionalnih kategorija kaznenog prava odnosi se na područja sustava UI-ja s većim stupnjem autonomije, prema kojima stremi trenutačni tehnološki razvoj i eksperimentiranje na mnogim poljima, što znači da će njihov značaj u bliskoj budućnosti nedvojbeno porasti.
16. Iz ove perspektive područje kaznene odgovornosti pravnih osoba moglo bi poslužiti kao koristan izvor, kao i područje odgovornosti za proizvode i odgovornosti za zaštitu zdravlja i sigurnosti na radnom mjestu.
17. U ovim pravno uređenim područjima, često usklađenima na europskoj razini, postojeća bi se načela, uz potrebne prilagodbe, mogla proširiti na propise o kaznenim djelima povezanim s UI-jem. Takva uređena područja zahtijevaju preventivnu procjenu rizika svojstvenih specifičnim aktivnostima koje se izvode, u kojima postoje granice dopuštenog rizika i povezane obveze prevencije i ograničavanja rizika (vidi 2. poglavlje), s posebnim obzirom na izvore opasnosti i štete.
18. Dužnosti djelovanja, posebno u slučaju upozorenja (crvenih zastavica; *red flags*), nametnute su relevantnim kategorijama osoba (ljudima), koje djeluju u skladu sa svojim kompetencijama, tj. korisnicima i osobama koje imaju položaj jamaca. Oni odmah moraju prilagoditi regulatorne i sigurnosne mjere primjerene svojoj djelatnosti, pa čak i zaustaviti je ako je potrebno.

19. Na temelju tih priznatih načela mogu se razraditi sljedeće preporuke za strukturiranje kaznene odgovornosti za štete povezane s UI-jem:

- a. Kaznena odgovornost fizičkih osoba. Ona se mora temeljiti na identifikaciji osobnih garantnih pozicija u odnosu na kompetencije i funkcije koje se obavljaju u korištenju sustava UI-ja. Kao prvo, treba uzeti u obzir di-onike koji sudjeluju u različitim fazama korištenja sustava UI-ja, sve do krajnjih korisnika. Kao drugo, treba razmotriti položaje visokog i srednjeg menadžmenta i službenika nadležnih za praćenje usklađenosti u složenim organizacijama. U svakom slučaju, treba formalizirati pozitivne obveze tehničke, organizacijske i kontrolne prirode.

Kaznena odgovornost za nehajna ponašanja mora biti u skladu s općim načelima kaznenog prava, odnosno s načelom osobne krivnje, budući da objektivna veza između uzročnog doprinosa ljudskog subjekta i počinjenja kaznenog djela od strane sustava UI-ja nije dovoljna s obzirom na to da je također potrebna predvidljivost i mogućnost izbjegavanja nezakonite radnje. Konkretno, kaznena odgovornost za nehaj – za propust djelovanja na način koji bi bio moguć – ne mora nužno biti vezana uz specifični i konkretni događaj ili činjenicu koja se dogodila koliko s pojmom „organizacijskog propusta“, koji se odnosi na način na koji je umjetni subjekt strukturiran i djeluje. Procjena rizika koji proizlaze iz aktivnosti sustava UI-ja mora uključivati i svijest o ishodima koji su konkretno „nepredvidivi“ u pojedinim slučajevima, što je temelj za obvezu pripreme adekvatnih i uvijek ažurnih mjera nadzora i sprječavanja za čije je ispunjenje odgovorna fizička osoba, koja je ujedno odgovorna kao vlasnik ili najviši predstavnik organizacije koja koristi sustav UI-ja u vlastitom interesu ili u vlastitu korist.

- b. Punitivna odgovornost pravnih osoba. S obzirom na to da velik dio sustava UI-ja proizvode ili koriste pravne osobe, nužno je smatrati ih odgovornima za djela počinjena od strane, putem ili protiv takvih sustava.

U tom pogledu, pod pretpostavkom uvođenja preciznih javnih standarda ponašanja i usklađenosti (usp. s 2. poglavljem), kažnjavanje pravne osobe, razmjerno kaznenim djelima počinjenima putem, od strane ili protiv sustava UI-ja i stupnju krivnje organizacije, može biti povezano s modelom odgovornosti temeljenim na organizacijskom propustu. Takav model odgovornosti dovodi do subjektivnog uračunavanja odgovornosti pravnoj osobi kao objektu kaznenog prijekora, za djela uzrokovana nedostatkom, manjkavošću ili neadekvatnošću organizacijskih i preventivnih mjera koje je potrebno provoditi i ažurirati na temelju procjena specifičnih rizika koji proizlaze iz aktivnosti koje su povjerene sustavima UI-ja i, u svakom slučaju, izvršenih od strane sustava UI-ja u interesu ili u korist te pravne osobe. Treba izbjegavati modele objektivne odgovornosti.

Treba promicati novi model autonomne kaznene odgovornosti pravnih osoba, koji se ne temelji na odgovornosti pojedinačne fizičke osobe, budući da se pravna osoba može smatrati odgovornom čak i ako fizička osoba koja je počinila štetu nije individualno kažnjiva zbog posebnih uvjeta ili okolnosti, ili u slučajevima kada nije identificirana. Doista, dovoljno je utvrditi počinjenje objektivno tipične i protupravne radnje u interesu ili u korist organizacije.

U nacionalnim pravnim sustavima koji odgovornost pravnih osoba temelje na zatvorenom popisu kaznenih djela preporučuje se proširenje takva popisa na kaznena djela koja se mogu počinuti putem, od strane ili protiv sustava UI-ja.

5. O preventivnim mjerama i kaznenim sankcijama koje se primjenjuju na fizičke i pravne osobe koje stoje „iza” sustava umjetne inteligencije

20. Sankcije koje se u različitim pravnim sustavima primjenjuju na fizičke osobe, uključujući i zatvorsku kaznu, odnosno na pravne osobe, čija je priroda katkada upravna, no u svakom slučaju punitivna, uključujući i novčane kazne i zabranu obavljanja djelatnosti kojom je djelo počinjeno, načelno trebaju odgovarati onima koje se primjenjuju za tu vrstu počinjenog djela, u skladu s načelima svakog pravnog sustava, odnosno načelima razmjernosti i individualizacije sankcija. Kad je riječ o pravnim osobama, te bi sankcije, uz novčane mjere, trebale uključivati i zabranu izmjene sustava usklađenosti i unutarnje kontrole pravne osobe, kao i mogućnost određivanja razdoblja javnog nadzora pravne osobe kako bi se osiguralo da ona zadovoljava postavljene standarde.
21. S obzirom na ozbiljnost štete povezane s UI-jem, preporučuje se usvajanje preventivnih mjera za učinkovitost u sprječavanju ili ublažavanju štetnih posljedica sustava UI-ja (kao što su, između ostalog, zapljena, oduzimanje predmeta, sudski nadzor, mjere zabrane).
22. Treba naglasiti važnu ulogu nematerijalnih sankcija, kao što su zabrana obavljanja određenih djelatnosti i oduzimanje predmeta. Konkretno oduzimanje predmeta omogućuje poduzimanje izravnih radnji protiv sustava UI-ja kojim ili pomoću kojega je počinjeno djelo, bez potrebe da se prizna kao pravna osoba ili da postoji mogućnost njezine kaznene odgovornosti (vidi poglavlje 1. gore).

6. Dopunski sustavi provedbe

23. S obzirom na problematične i predvidljive teškoće u provedbi učinkovitog sustava kaznene odgovornosti za fizičke i pravne osobe koje su „iza” sustava UI-ja za kaznena djela počinjena od strane, putem ili protiv njih preporučuje se donošenje dopunskih sustava provedbe.
24. Takav sustav provedbe može uključivati upravna odobrenja i certifikacije, kao i građanskopravna sredstva zaštite.
25. Ostale alternative kaznenom progonu i/ili sankcioniranju mogu uključivati modele usklađenosti, intervencije restorativne pravde i sporazume sa žrtvama i nadležnim javnim tijelima.

II. SEKCIJA: KAZNENO PRAVO I KRIMINALIZACIJA SUOČENA S IZAZOVIMA UMJETNE INTELIGENCIJE **

Rezolucija je usvojena na Kolokviju „Stara i nova kaznena djela:
sustavi UI-ja kao instrumenti i žrtve“, Bukurešt, 14.–16. lipnja 2023.

Preambula

S obzirom na to

- da je umjetna inteligencija (dalje: UI), jedno od posljednjih dostignuća digitalne revolucije, već dosegla značajnu razinu razvoja u ovom trećem desetljeću 21. stoljeća te se već naširoko koristi u mnogim društvenim sektorima;
- da, iako ne postoji apsolutna suglasnost oko definicije te tehnologije, postoji implicitni konsenzus da ona obuhvaća mnoštvo kompjutoriziranih sustava koji su, putem prikupljanja, obrade i analize podataka u njihovu kontekstu, sposobni djelovati autonomno i/ili pomagati u donošenju odluka radi postizanja specifičnih ciljeva;
- da transformativni potencijal te tehnologije utječe na više područja i društvenih sfera donoseći važne prednosti i prilike;
- da UI također stvara rizike i štete za pojedinačne i kolektivne interese;

Primjećujući

- da se višestruke odluke koje su tradicionalno donosili ili na koje su utjecali ljudi počinju automatizirati korištenjem tih tehnologija, što utječe na različita područja i interese;
- da je u područjima kao što su autonomna vozila, zdravstvene usluge, financijska tržišta, mediji te u drugim sektorima korištenje te tehnologije nezaustavljivo i da se budućnost s njom čini neizbježnom;
- da obećanje da je UI istovremeno učinkovit i objektivan dovodi do razvoja tih tehnologija bez procjene njihove stvarne potrebe ili razmatranja rizika koje mogu stvoriti;

Imajući na umu

- da je nedavni razvoj velikih jezičnih modela i drugih sustava UI-ja, kao što su strojno učenje i duboko učenje, istaknuo potrebu za regulacijom, uključujući i sigurnosne protokole, kako bi se kontrolirao razvoj tih tehnologija s obzirom na njihove učinke i rizike;
- da razvoj sustava UI-ja, osobito treniranje njegovih algoritama, zahtijeva korištenje i akumulaciju podataka i velike količine informacija, što predstavlja rizik koji je potrebno zasebno razmotriti;

** Rezoluciju je prevela Darija Željko Mrljak, mag. iur., LL. M., asistentica na Katedri za kazneno procesno pravo Pravnog fakulteta Sveučilišta u Zagrebu, pri čemu se kao pomoćnim alatom koristila platformama UI *ChatGPT* i *Google Translate*. Prijevod je temeljito provjeren, ispravljen, kontekstualiziran i usklađen sa standardnom pravnom terminologijom.

Ističući

- sve veću zabrinutost zbog štete koju zlonamjerno ili nepažljivo korištenje UI-ja može prouzročiti u područjima u kojima UI već ostvaruje snažnu prisutnost;
- da je u mnogim državama korištenje sustava UI-ja prouzrokovalo štetu relevantnim interesima kao što su život, zdravlje i privatnost, između ostalih;

Priznajući

- da će pojava novih kaznenih djela, kao i novih interesa koji zaslužuju kaznenu zaštitu, potaknuti države na prilagodbu kaznenih zakona u vezi s UI-jem;
- potrebu za analizom je li pravni odgovor država na izazove UI-ja dovoljan ili ga je potrebno reformirati i prilagoditi, bilo kroz specifične izmjene bilo uspostavljanjem novih oblika kriminalizacije;

Uzimajući u obzir

- Preporuku Vijeća o umjetnoj inteligenciji, koju je OEES usvojio 22. svibnja 2019.; Prijedlog Uredbe Europskog parlamenta i Vijeća o utvrđivanju usklađenih pravila o umjetnoj inteligenciji (Akt o umjetnoj inteligenciji) i izmjenu određenih zakonodavnih akata Unije koje je Europska komisija predstavila 21. travnja 2021.; Rezoluciju Europskog parlamenta od 6. listopada 2021. o umjetnoj inteligenciji u kaznenom pravu i njezinoj primjeni od strane policijskih i pravosudnih tijela u kaznenim stvarima (2020/2016(INI)); Etičke smjernice za pouzdanu umjetnu inteligenciju, koje je 8. travnja 2019. predstavila Stručna skupina na visokoj razini osnovana od Europske komisije; te Preporuku CM/Rec (2020) 1 Odbora ministara državama članicama o utjecaju algoritamskih sustava na ljudska prava (usvojenu od Odbora ministara 8. travnja 2020. na 1373. sastanku ministarskih zamjenika);
- rasprave i rezolucije prijašnjih međunarodnih kongresa kaznenog prava, posebice rezolucije XIV. Kongresa o kaznenom pravu i suvremenim biomedicinskim tehnikama te pravnim i praktičnim problemima koji proizlaze iz razlika između kaznenog prava i upravnog kaznenog prava; rezolucije XIX. Međunarodnog kongresa kaznenog prava – Informacijsko društvo i kazneno pravo; te nacрте rezolucija XXI. Međunarodnog kongresa kaznenog prava I. sekcije (Kazneno pravo – opći dio) i III. sekcije (UI i upravljanje kaznenim pravosuđem);

Sudionici II. sekcije Kolokvija u Bukureštu, održanog od 14. do 16. lipnja 2023., usvojili su sljedeće:

REZOLUCIJE

I. Regulacija i provedba propisa o UI-ju

1. S razvojem i širenjem UI-ja pojavili su se, a vjerojatno će se i povećati, štete i rizici za individualne i kolektivne interese. To nameće potrebu za odgovarajućom regulacijom proizvodnje, implementacije i uporabe sustava UI-ja. Odgovor na te izazove treba razmotriti različite perspektive i sve dostupne regulatorne alate, bilo da su javni bilo privatni, uzimajući u obzir različitu prirodu i funkcionalnost svakog od njih.
2. Globalni utjecaj UI-ja zahtijeva međunarodni odgovor kako bi se učinkovito zaštitili individualni i kolektivni interesi koji su ugroženi. Države bi trebale uzeti u obzir međunarodne standarde prilikom donošenja i provedbe nacionalnih propisa.
3. Kao što se dogodilo s drugim tehnološkim i društveno-ekonomskim inovacijama, primjerice s internetom ili novim otkrićima u području uređivanja gena (*gene editing*) i neuroznanosti, nadiranje UI-ja nužno zahtijeva preispitivanje općih aspekata kaznenopravnog sustava i, posebice, postojećeg kataloga kaznenih djela u kaznenim zakonima kako bi se utvrdilo jesu li državni propisi prikladni za izazove koje postavlja uporaba te tehnologije.
4. Rasprave o transformaciji kaznenog prava kao odgovoru na utjecaj UI-ja i o ulozi koju kaznenopravni sustav može igrati u odnosu na te nove tehnologije ne smiju biti izolirane od međunarodnih etičkih rasprava koje se vode o razvoju UI-ja niti mogu zanemariti rasprave o regulaciji UI-ja u drugim granama pravnog sustava. Uzimajući u obzir taj široki okvir, zakonodavci bi trebali razmisliti o specifičnoj ulozi koju kazneno pravo ima u sprječavanju šteta koje proizlaze iz uporabe UI-ja.
5. Zakonodavci će morati reformirati postojeća kaznena djela kada UI promijeni rizik za postojeće interese ili stvori nove načine počinjenja kaznenih djela koji nisu obuhvaćeni postojećim zakonodavstvom. Također, morat će se uvesti nova kaznena djela kad razvoj te tehnologije dovede do pojave novih individualnih i kolektivnih interesa koji zahtijevaju zaštitu, a koje postojeće zakonodavstvo ne pokriva.

II. Kriminalizacija i zaštita interesa povezanih s UI-jem

6. Razvoj UI-ja može dovesti do pojave novih interesa koji su vrijedni zaštite. Uz to, sustavi UI-ja mogu utjecati na dimenziju i relevantnost interesa koji se trenutno ne smatraju vrijednima kaznenopravne zaštite. Kada kazneni zakoni ne pružaju odgovarajući odgovor za zaštitu takvih interesa, potrebno je uvesti nova kaznena djela koja proporcionalno kažnjavaju ponašanja koja im štete. To se treba učiniti samo kada ne postoje sredstva manje štetna od kaznenog prava za učinkovitu zaštitu tih interesa.
7. Kada transformacija UI-ja dovede do nastajanja novih interesa koji su u suštini slični drugima koji su tradicionalno smatrani vrijednima zaštite, ne treba uvoditi nova kaznena djela. Umjesto toga preporučuje se prilagoditi tumačenje postojećih kaznenih djela, pod uvjetom da strogo poštivanje načela zakonitosti to dopušta.

8. Zakonodavci će morati odlučiti hoće li razvoj ove tehnologije zahtijevati specifičnu kaznenopravnu zaštitu individualnih ili kolektivnih interesa povezanih sa samom tehnologijom UI-ja. Iako je još prerano da se utvrdi hoće li takva potreba nastati, to bi mogao biti slučaj za podatke na kojima se temelje algoritmi; za funkcionalnost samih sustava UI-ja u određenim slučajevima; za kolektivne interese vezane uz sigurnost i pouzdanost dizajna i primjene UI-ja; pa čak i za interese povezane s robotima.
9. Neki sustavi UI-ja, poput onih koji se koriste u kritičnim infrastrukturama, ključni su za osiguravanje već zaštićenih interesa. U mjeri u kojoj Budimpeštanska konvencija predviđa kriminalizaciju napada na računalne sustave, a UI se može smatrati takvim sustavom, uvođenje novih kaznenih djela možda neće biti potrebno. Kako bi se izbjegle sumnje u tumačenju, preporučljivo bi bilo reformirati neka kaznena djela kako bi se sustavi UI-ja uveli kao vrsta računalnih sustava.
10. Ako se ne dokaže postojanje nekih drugih interesa, te uzimajući u obzir trenutni stupanj razvoja tih tehnologija, UI i robotski sustavi ne zaslužuju različitu zaštitu od ostalih računalnih sustava uzimajući u obzir njihovu ekonomsku ili funkcionalnu vrijednost.

III. Osnove legitimizacije i tehnike kriminalizacije

11. Kazneno pravo ne smije imati vodeću ulogu u regulaciji UI-ja. S obzirom na njegovu prirodu kao osobito prisilnog instrumenta provedbe, ono bi trebalo intervenirati kao posljednja mjera i biti ograničeno na suzbijanje najozbiljnijih i najštetnijih djela.
12. Zakonodavci ne smiju uvoditi nova kaznena djela samo na temelju činjenice da je korišten UI. Mnoga kaznena djela mogu se počiniti korištenjem sustava UI-ja kao sredstva za počinjenje kažnjivog ponašanja. Tek kada djela počinjena pomoću sustava UI-ja poprime drugačije značenje u smislu štetnosti ili rizika, bit će potrebno uvesti nova kaznena djela.
13. Automatizacija procesa donošenja odluka na temelju podataka, što omogućuje UI, premješta ključni trenutak ljudskog djelovanja u faze dizajna i implementacije algoritama koje su daleko od samog nastanka štete. Odgovornost pojedinaca i pravnih osoba uključenih u taj proces trebala bi stoga biti usmjerena na navedene faze. To će se učiniti uz uzimanje u obzir postojećih pravnih obveza utvrđenih u drugim granama pravnog sustava.
14. Kaznenopravni sustavi osmišljeni su kako bi imali odvrćajući učinak na vjerojatne počinitelje sprečavajući ih u poduzimanju kaznenih radnji. Ako je ključni trenutak u pogledu rizika u vezi s umjetnom inteligencijom faza dizajna i implementacije, trebalo bi razmotriti uvođenje kaznenih djela koja imaju za cilj odvrćanje ponašanja u tim fazama. To se može postići anticipirajući zaštitu kroz kaznena djela ugrožavanja koja kažnjavaju neispunjavanje određenih obveza u vezi sa specifičnim interesima koji su vrijedni zaštite. Također, slično kao što je propisano za kaznenu odgovornost pravnih osoba, mogle bi se uspostaviti posebne regulatorne obveze u vezi s dizajnom i implementacijom sustava UI-ja, čije kršenje može dovesti do kaznene odgovornosti.

15. Kaznena djela ugrožavanja povezana s dizajnom i implementacijom sustava UI-ja koja uzrokuju rizike trebala bi biti uvedena kada sankcionirane radnje predstavljaju značajnu prijetnju za zaštićene interese. Dodatno, pravne posljedice povezane s tim kaznenim djelima trebaju biti proporcionalne razini rizika i značaju ugroženog interesa. Zbog složenosti dizajna sustava UI-ja i različitih pristupa regulaciji tih alata kaznena djela ugrožavanja ne bi se smjela uvoditi prije negoli se uzme u obzir razvoj samoregulacije ili upravnih propisa o kontroli i sigurnosti UI-ja u svakom pravnom sustavu. Ti propisi trebali bi služiti za identificiranje relevantnih rizičnih radnji koje bi mogle zaslužiti kazneni progon.
16. Moguće je uvesti nova kaznena djela za kažnjavanje zloporabe i transformaciju postojećih zakonitih sustava UI-ja kada promjenom dizajna ili svrhe UI-ja nastaju novi rizici.
17. U onim pravnim sustavima gdje se nemarno ponašanje kažnjava samo kada je to izričito predviđeno (*numerus clausus*) mogla bi biti potrebna reforma kaznenih zakona. S obzirom na složen dizajn sustava umjetne inteligencije i sudjelovanje više strana u njegovu životnom ciklusu, u većini slučajeva bit će iznimno teško dokazati svijest o mogućem štetnom ishodu u trenutku dizajna. Stoga se mogu uvesti kaznena djela iz nehaja na temelju kršenja standarda dužne pažnje ako to zahtijeva zaštita pogođenih interesa.
18. Budući da su sustavi UI-ja dinamični i njihova izvedba ovisi o uvođenju ili prikupljanju podataka koji mijenjaju njihove rezultate, procesi upravljanja rizicima uspostavljeni u drugim granama pravnog sustava mogli bi biti primjenjivi tijekom cijelog životnog ciklusa UI-ja. Stoga kazneni zakoni mogu, ako je potrebno, obuhvatiti kršenja pravila vezana uz nedostatak odgovarajućeg nadzora i praćenja sustava UI-ja, što su obveze koje mogu utjecati na različite subjekte uključene u cijeli životni ciklus UI-ja.
19. Budući da su tehnologija UI-ja i njezine primjene podesive, kaznenopravni sustavi mogli bi se prilagoditi kako bi se uskladila proporcionalnost kazni prema težini štete koju UI može uzrokovati. Ipak, zakonodavci ne bi smjeli uvesti otegotne okolnosti samo zato što je kazneno djelo počinjeno korištenjem UI-ja. Samo ako postojeće otegotne okolnosti ne mogu obuhvatiti težinu štete uzrokovane korištenjem UI-ja, s obzirom na važnost pogođenog interesa, trebale bi se razmotriti nove otegotne okolnosti. To bi se uvijek trebalo činiti u skladu s načelom proporcionalnosti.

IV. Kriminalizacija i zaštita specifičnih interesa od rizika koje stvara UI

20. Budući da kazneni zakoni obično ne predviđaju specifična sredstva za počinjenje kaznenih djela protiv života i zdravlja, ne čini se potrebnim reformirati navedena kaznena djela kako bi se zaštitili ti interesi kada je sustav UI-ja korišten kao sredstvo počinjenja. Također, ne čini se nužnim mijenjati sustav stupnjevanja odgovornosti. Ipak, u specifičnim područjima, poput autonomne vožnje, kazneno pravo morat će biti pažljivo prema promjenama u regulatornom okruženju, koje će biti temelj za utvrđivanje što se smatra „dopuštenim rizikom“ za određivanje odgovornosti.

21. Ako autonomna vožnja postane široko rasprostranjena, kaznena djela povezana sa sigurnošću na cestama mogla bi doživjeti značajne promjene, uključujući i nova kaznena djela vezana uz nove rizične oblike ponašanja u vezi sa životom i sigurnošću na cestama koja nisu usmjerena isključivo na ljudsku vožnju.
22. Na jednak način na koji je revolucija u genskoj modifikaciji dovela do pojave kaznenih djela koja kažnjavaju gensku manipulaciju s kapacitetom za masovno uništenje tehnološki razvoj mogao bi uskoro učiniti nužnim kriminalizaciju stvaranja, razvoja i korištenja alata UI-ja s visokim destruktivnim kapacitetom, kao što su neka autonomna oružja, dronovi ili roboti, koji bi mogli biti iznimno štetni, osobito ako se izgubi ljudska kontrola nad njima.
23. Sustavi UI-ja prikupljaju velike količine informacija za obavljanje svojih zadataka i oslanjaju se na njih, čime stvaraju nove prijetnje klasičnim interesima. S obzirom na takav razvoj tehnologije UI-ja potrebno je preispitati kaznena djela povezana s privatnošću i drugim osobnim interesima te razmotriti promjenu shvaćanja privatnosti kao isključivo individualnog dobra, uzimajući u obzir kolektivnu dimenziju navedenog interesa.
24. Kriminalizacija radnji koje uključuju nezakonito prikupljanje osobnih podataka ne bi trebala biti vezana samo uz zaštitu interesa poput privatnosti. Korištenje UI-ja u kibernetičkom prostoru može otvoriti vrata masovnom prikupljanju podataka za počinjenje kibernetičkih kaznenih djela koja ugrožavaju interese poput imovine. Države bi trebale razmotriti je li potrebno uvesti kaznena djela kojima bi se sankcioniralo nezakonito masovno prikupljanje podataka, kao i slične pripremne radnje za teška kaznena djela, kada to uzrokuje konkretan rizik za navedene interese i samo ako ne postoji drugi manje restriktivan pravni alat.
25. Dostupnost slika i osobnih podataka u kibernetičkom prostoru u kombinaciji s potencijalom generativnog UI-ja da transformira slike, videozapise i zvuk može ugroziti interese poput ugleda, časti ili seksualne slobode. Potrebno je preispitati omogućuju li trenutačni kazneni zakoni kažnjavanje ponašanja štetnih za ljudsko dostojanstvo, ugled i seksualnu slobodu, poput distribucije „Deep Fakeova”, uključujući i one sa seksualnim sadržajem, ili dječje pornografije.
26. Generativni UI, poput velikih jezičnih modela i sličnih alata, može olakšati prijevaru, prijetnje i prisilu utječući na različite faze formiranja volje te ugroziti interese vrijedne zaštite. Ipak, ne čini se prikladnim uvesti nova kaznena djela jer trenutačni kazneni zakoni obuhvaćaju štetnije radnje te bi se drugi načini kontrole tih alata trebali koristiti za sprječavanje manje rizičnih ponašanja.
27. Popularizacija algoritama za upravljanje rizicima u područjima kao što su zdravstvena zaštita, zapošljavanje radnika, pravosuđe, krediti i zajmovi te u mnogim drugima otkrila je postojanje diskriminatornih pristranosti u nekim odlukama koje donose sustavi UI-ja. Osim kaznenih djela kojima se već mogu kažnjavati neke osobito ozbiljne diskriminatorne odluke, druge su grane pravnog sustava, poput građanskog ili upravnog prava, prikladnije za sprječavanje problema algoritamske diskriminacije.

28. Korištenje UI-ja u kibernetičkom prostoru može olakšati i pojačati postojeće napade na imovinu i druge interese vrijedne zaštite. Međutim, s obzirom na regulaciju kibernetičkih prijevara i drugih kibernetičkih kaznenih djela protiv imovine neće biti potrebno, barem u kratkom roku, prilagođavati ta kaznena djela kako bi se obuhvatila kaznena djela počinjena korištenjem sustava UI-ja.
29. Neki kazneni zakoni koji kažnjavaju proizvodnju, prodaju, nabavu radi korištenja, uvoz, distribuciju ili druge oblike stavljanja na raspolaganje uređaja ponajprije izrađenih ili prilagođenih kako bi omogućili počinjenje bilo kojeg od kaznenih djela utvrđenih u člancima 2. do 5. Budimpeštanske konvencije već mogu kazniti stvaranje, razvoj i prodaju sustava UI-ja izrađenih ili prilagođenih za te kaznene svrhe. Prema tome, ako se UI smatra uređajem, uključujući i računalni program, u skladu s člankom 6. Budimpeštanske konvencije nije nužno uvoditi nova kaznena djela koja bi anticipirala kazneni odgovor u tom području.
30. U društveno-ekonomskom i financijskom području već se izvještava o naglom širenju sustava algoritamskog donošenja odluka i upotrebi UI-ja za trgovanje. Rizik da zlonamjerna ili nemarna uporaba sustava UI-ja ozbiljno ugrozi tržišta jasan je, ali bi se interesi koji su u pitanju mogli bolje zaštititi preventivnim regulatornim mjerama ekonomske, upravne i komercijalne prirode nego uvođenjem novih kaznenih djela, osim onih već postojećih za kažnjavanje insajderskog trgovanja i sličnog ponašanja.
Kada se sustavi UI-ja koriste za manipulaciju tržištima, potrebno je revidirati kazneno pravo kako bi se pružio proporcionalan odgovor.
31. Zabrinutost zbog utjecaja dezinformacija, prvo nakon nekih izbornih procesa i zatim zbog infodemije tijekom krize izazvane *Covidom-19*, potaknula je mnoge države da uvedu kaznena djela kojima se kažnjava takvo ponašanje. Činjenica da UI može povećati utjecaj dezinformacija bilo automatizacijom njihove diseminacije bilo korištenjem sofisticiranih tehnologija za manipulaciju videozapisima, zvukom, slikama i tekstom mogla bi dodatno podržati taj trend. Kriminalizacija dezinformacija bit će opravdana samo radi zaštite temeljnih interesa demokratskih društava i pod uvjetom da ta kriminalizacija ne ugrožava slobodu izražavanja.

III. SEKCIJA: UMJETNA INTELIGENCIJA I KAZNENO PRAVOSUĐE: „PROGNOSTIČKI RAD POLICIJE“, „PROGNOSTIČKO PRAVOSUĐE“ I DOKAZI ***

Rezolucija je usvojena na Kolokviju „Umjetna inteligencija i kazneno pravosuđe: prognostički rad policije i prognostičko pravosuđe“, Buenos Aires, 28.–31. ožujka 2023.

Preambula

Svjesni da se umjetna inteligencija (UI) ubrzano razvija u suvremenom društvu u različitim dijelovima svijeta. Već sveprisutan u životima ljudi u nekim državama, UI bi u budućnosti mogao postati dio svakodnevnog života velikog dijela svjetske populacije.

Napominjući da kao tehnološka inovacija UI potiče potrošače na kupnju novih proizvoda, čime pomaže rastu globalne ekonomije. Stoga UI ima nezanemarljivu ulogu u održavanju, pa čak i širenju liberalne tržišne ekonomije i kapitalističkog gospodarskog sustava.

Napominjući da su tvrtke koje stvaraju i plasiraju UI često smještene u razvijenim zemljama globalnog Sjevera te često nastoje otvoriti tržišta diljem svijeta.

Napominjući da digitalni jaz produbljuje društvene nejednakosti među ljudima. „Jaz u UI-ju“ mogao bi biti sljedeći fenomen na horizontu.

Razmatrajući da se UI može definirati kao skup teorija i tehnika koje se koriste za stvaranje strojeva sposobnih simulirati ljudsku inteligenciju.¹ Kao znanstvena disciplina, UI predstavlja spoj statističke i algoritamske matematike, računalnih znanosti i kognitivnih znanosti. Simbolički UI (*Symbolic AI*) temelji se na pravilima logike, dok konektivistički UI (*connectionist AI*) koristi umjetne neuronske mreže.

Razmatrajući da je strojno učenje primjer konektivističkog UI-ja, kao i duboko učenje (*deep learning*), koje je podskup strojnog učenja (*machine learning*), koje koristi više slojeva međusobno povezanih umjetnih neurona. Kako se broj neuronskih slojeva koji omogućuju autonomno učenje povećava, povećava se i tehnološka složenost sustava čineći sustav učinkovitijim, a njegove izračune manje objašnjivim i sljedivim (duboko učenje).

Razmatrajući da strojno učenje može učiniti tehnologiju iznimno moćnom, ali njezin proces donošenja odluka može biti toliko složen da nalikuje „crnoj kutiji“.

Napominjući da su mnogi sustavi UI-ja koji se koriste u kontekstu sprečavanja, istraživanja, otkrivanja i kažnjavanja kaznenih djela sustavi strojnog učenja. Koristeći

*** Rezoluciju je prevela dr. sc. Zlata Đurđević, redovita profesorica Pravnog fakulteta Sveučilišta u Zagrebu, pri čemu se kao pomoćnim alatom koristila platformama UI-ja ChatGPT i Google Translate. Prijevod je temeljito provjeren, ispravljen, kontekstualiziran i usklađen sa standardnom pravnom terminologijom.

¹ https://www.larousse.fr/encyclopedie/divers/intelligence_artificielle/187257

samoučeće (*self-learning*) algoritme, oni izvode složene izračune vjerojatnosti u nanosekundama. Kako bi postigli zadane ciljeve, obrađuju ogromne količine podataka i troše veliku količinu energije. Neki od njih, poput sustava za prepoznavanje lica, oslanjaju se na duboko učenje.

Utvrđujući da se u kaznenom pravosuđu sustavi UI-ja koriste za sprečavanje ili otkrivanje kaznenih djela na temelju procjene rizika („prognostički rad policije“).

Utvrđujući da se sustavi UI-ja također koriste za pomoć tužiteljima i sucima u donošenju odluka. Točnije, izraz „prognostičko pravosuđe“ odnosi se na (i) predviđanje nečijeg ponašanja, npr. procjenu rizika bijega u istražnom postupku ili (ponovnog) počinjenja kaznenog djela, kako bi se o njemu ili njoj donijele odluke kao što su istražni zatvor, izricanje kazne, uvjetni otpust i uvjetna osuda (aktuarsko pravosuđe (*actuarial justice*), koje danas može biti podržano UI-jem); i (ii) korištenje UI-ja za ultrabrzu statističku analizu ranijih odluka donesenih u sličnim slučajevima i relevantnih pravnih te regulatornih odredbi (kvantitativna pravna analiza ili *LegalTech*).

Smatrajući da je kvantitativna pravna analiza revolucionarna u smislu da je matematički izračun zamišljen kao podrška ili čak zamjena pravnog razmišljanja.

Primjećujući da je riječ „prognostički“, koja se koristi u frazama „prognostički policijski rad“ (*predictive policing*) i „prognostičko pravosuđe“ (*predictive justice*), zbuñujuća jer sustavi UI-ja izračunavaju vjerojatnosti, ali ne predviđaju budućnost; te vjerojatnosti temelje se na korelacijama, a ne na uzročno-posljedičnim vezama. Ti izračuni ipak imaju performativni učinak na ljude, odnosno mogu ih potaknuti da donose odluke u skladu s njihovim rezultatima. Općenito govoreći, znanstveni temelji UI-ja potiču njegove korisnike da vjeruju i slijede vjerojatnosti koje je izračunao sustav UI-ja jer je „automatizacijska pristranost“ (*automation bias*) veća kada sustav sadrži u određenoj mjeri znanstvenu auru.

Utvrđujući da sustavi UI-ja pridonose inovacijama u potrazi za dokazima. Brzo mogu analizirati velike podatkovne skupove (*big data*) i izvući informacije koje mogu biti korisne istražiteljima. Sustavi UI-ja mogu uspostaviti korelacije između informacija koje su nevidljive ljudskom oku. Dijagrami analize kriminala koje proizvode mogu biti vrlo vrijedni, detaljni podaci za istražitelje.

Utvrđujući da sustavi UI-ja mogu proizvesti informacije koje se koriste kao dokaz u kaznenim postupcima. Konkretno, sustavi UI-ja mogu pružiti forenzičke informacije uspoređivanjem biometrijskih obilježja (npr. slika lica u sustavu za prepoznavanje lica), zvučnih frekvencija različitih glasova (prepoznavanje glasa) i fragmenata DNK (probabilističko genotipiziranje).

Utvrđujući, naposljetku, da roboti potpomognuti UI-jem i „pametni“ predmeti u različitim zanimanjima i svakodnevnom životu mogu nehotice proizvesti tragove ili dokaze koji bi mogli biti korisni u utvrđivanju činjenica u kaznenom predmetu.

Utvrđujući da, unatoč značajnom napretku posljednjih godina, sustavi UI-ja nisu potpuno pouzdani. Pogreške mogu biti rezultat loše kvalitete podataka koji se koriste ili načina na koji je algoritam programiran ili prisutnosti lažnih pozitivnih/negativnih korelacija. Stoga vjerojatnosti koje proizvede sustav UI-ja mogu biti netočne.

Utvrđujući da rezultati koje proizvode sustavi UI-ja nisu uvijek potpuno neutralni. Točnost vjerojatnosti izračunatih od strane sustava UI-ja ovisi ne samo o kvaliteti prikupljenih i obrađenih podataka, koji mogu odražavati pristranost, već i o tome kako su sustavi učili (nenadzirano naspram nadziranog učenja). Budući da reproduciraju ljudske odluke, samoučeći algoritmi podložni su ljudskim slabostima. Jedan su od rezultata ksenofobni,² rasistički, mizogini itd. algoritmi.

Utvrđujući da sustavi UI-ja postavljaju probleme transparentnosti. Takozvana „crna kutija UI-ja“ toliko je neprozirna da čak ni stručnjaci ne mogu odrediti kako dolazi do svojih rezultata. Čak ni znanstveni stručnjaci ne mogu u potpunosti objasniti rasuđivanje (*reasoning*) sustava sudu.

Razmatrajući da sustave UI-ja koji se koriste u području kaznenog pravosuđa može razvijati privatni sektor. Takvi su sustavi proizvodi koji se prodaju i moraju biti profitabilni. Tvrtke koje ih razvijaju uglavnom se pozivaju na zaštitu poslovne tajne kako bi odbile otkriti izvorni kod svog algoritma, bez kojeg se funkcioniranje sustava ne može ispravno analizirati.

Uočavajući da nije sve što je tehnološki moguće ujedno i društveno poželjno. U demokraciji se političke odluke koje utječu na prevenciju, otkrivanje, istragu i kažnjavanje kaznenih djela moraju odražavati u zakonu ili normi iste obvezujuće snage.

Ponavljajući da ljudska prava moraju biti u potpunosti zaštićena tijekom prevencije, otkrivanja, istrage i kažnjavanja kaznenih djela, uključujući i situacije kada se u tom kontekstu koriste tehnološke inovacije. Budući da UI često otvara pitanja vezana uz privatnost i zaštitu osobnih podataka, kao i zakon o nediskriminaciji, svi zakoni koji štite ljudska bića, posebno njihovu slobodu i dostojanstvo, kao i sve garancije pravičnog suđenja, uključujući i pretpostavku nedužnosti, potencijalno su ugroženi korištenjem tehnologija koje simuliraju ljudsku inteligenciju.

Uzimajući u obzir da nacionalni zakoni i međunarodne i/ili regionalne pravne norme mogu definirati uvjete pod kojima se tehnološke inovacije povezane s UI-jem mogu koristiti u kaznenom pravosuđu.

Ponavljajući da etički standardi na koje se često poziva privatni sektor nemaju istu obvezujuću snagu kao zakon.

Svjesni:

- Preporuke Vijeća o umjetnoj inteligenciji, Organizacija za ekonomsku suradnju i razvoj, 22. svibnja 2019., C/MIN(2019)3/FINAL;
- Preporuke o etici umjetne inteligencije, Organizacija Ujedinjenih naroda za obrazovanje, znanost i kulturu, 23. studenoga 2021., SHS/BIO/PI/2021/1;
- Europske etičke povelje o upotrebi umjetne inteligencije u pravosudnim sustavima i njihovu okruženju, Europska komisija za učinkovitost pravosuđa, Strasbourg, 3.–4. prosinca 2018.;
- Rezolucije Europskog parlamenta od 6. listopada 2021. o umjetnoj inteligenciji i kaznenom pravu te njenoj upotrebi od strane policije i pravosudnih vlasti u kaznenim stvarima, dokument 2020, 2016 (INI);
- Rezolucije XIX. Međunarodnog kongresa za kazneno pravo: Informacijsko društvo i kazneno pravo, Rio de Janeiro, 2014.

² <https://www.amnesty.org/en/documents/eur35/4686/2021/en/>

REZOLUCIJE

1. Upotreba sustava UI-ja od strane javnih vlasti za pomoć u sprečavanju, otkrivanju ili istraživanju kaznenih djela mora biti unaprijed odobrena zakonom ili normom jednake obvezujuće snage.
2. Države moraju osigurati da su odluke vlasti da se usmjere na sprečavanje, otkrivanje ili istraživanje određene vrste kaznenog djela temeljene na političkim i demokratski donesenim kriterijima, a ne na pretpostavci da će upotreba tehnologije UI-ja olakšati sprečavanje, otkrivanje ili istraživanje te vrste kaznenog djela.
3. Kako bi zaštitile legitimnost aktivnosti javnih vlasti u sprečavanju, otkrivanju i istraživanju kaznenih djela, države koje žele koristiti sustave UI-ja moraju odabrati sustave čije je funkcioniranje u potpunosti transparentno, objašnjivo i sljedivo (bijela kutija UI-ja). Moraju osigurati da se ne mogu postaviti prigovori zbog intelektualnog vlasništva kada se traži transparentnost i trebaju dati prednost javno dostupnim sustavima otvorenih izvora (*open-source systems*).
4. Zakoni ili norme jednake pravne snage koji se odnose na korištenje sustava UI-ja u sprečavanju, otkrivanju i istraživanju kaznenih djela moraju zahtijevati da ti sustavi imaju visok stupanj tehnološke pouzdanosti. Dovoljno precizna regulacija koja zahtijeva odgovarajuće provjere i evaluacije, koje su vanjske i neovisne o razvojnom timu ili pružatelju sustava UI-ja, mora u najvećoj mogućoj mjeri ograničiti rizik od pristranosti ili bilo kojeg oblika diskriminacije u strojnom učenju, grešaka u kodiranju i drugih tehnoloških kvarova.
5. Zakoni ili norme jednake pravne snage moraju zahtijevati da sustavi UI-ja korišteni za pomoć u sprečavanju, otkrivanju i istraživanju kaznenih djela budu potpuno dostupni, provjerljivi i podložni reviziji od strane vlasti koje ih koriste te vlasti koje su odgovorne za provjere i evaluacije.
6. Zakoni ili norme jednake pravne snage koji dopuštaju korištenje sustava UI-ja za pomoć u sprečavanju, otkrivanju i istraživanju kaznenih djela moraju zahtijevati da podaci za obučavanje sustava budu visokokvalitetni i reprezentativni. Što se tiče podataka iz policijskih ili sudskih spisa, zakoni ili norme jednake pravne snage moraju uspostaviti sustav koji osigurava da su ti podaci točni i ažurirani te da njihova upotreba ne krši pretpostavku nedužnosti. Pretpostavka nedužnosti strogo zabranjuje zadržavanje i upotrebu podataka prikupljenih kao rezultat prognostičke procjene kada naknadno nije donesena presuda o krivnji, osim ako ti podaci nisu relevantni za drugog osumnjičenika. Što se tiče drugih podataka, osobito podataka dostupnih na društvenim mrežama, zakoni ili norme jednake pravne snage moraju zahtijevati poštivanje prava na privatnost i zakona o zaštiti osobnih podataka prilikom korištenja tih podataka. Potrebno je provesti odgovarajuće provjere, neovisne o policijskim i pravosudnim institucijama. Općenito, zakoni ili norme jednake pravne snage moraju biti iznimno zahtjevniji u pogledu provjere pouzdanosti svih podataka koje koriste sustavi UI-ja u vezi s otkrivanjem, sprječavanjem i istraživanjem kaznenih djela.

7. U pravilu, zakoni ili norme jednake pravne snage moraju zahtijevati da, prije nego što se sustav UI-ja temeljen na algoritmima koji uče samostalno može koristiti u sprečavanju, otkrivanju ili istraživanju kaznenih djela, ti algoritmi moraju biti razvijeni, obučeni, testirani i primijenjeni pod ljudskim nadzorom (učenje stroja s ljudskom kontrolom (*human-in-the-loop machine learning*)).

Ovi zakoni i norme jednake pravne snage moraju zahtijevati ljudsku procjenu prije poduzimanja bilo koje radnje za sprečavanje, otkrivanje ili istraživanje kaznenih djela na temelju vjerojatnosti koje je izračunao sustav UI-ja.

8. Države i tijela za provođenje zakona moraju osigurati da njihovo osoblje koje koristi UI za sprečavanje, otkrivanje ili istraživanje kaznenih djela dobije praktičnu obuku o pravilnoj uporabi relevantnog sustava UI-ja, kao i obuku o riziku od pogreške i pristranosti. Moraju osigurati da to osoblje ima temeljito znanje o opasnostima koje UI može predstavljati za ljudska prava.

9. Međunarodna, regionalna, nacionalna ili lokalna tijela moraju uspostaviti neovisna tijela koja certificiraju kvalitetu sustava UI-ja namijenjenih sprečavanju, otkrivanju ili istraživanju kaznenih djela. Tehnologija UI-ja kojom se ne može upravljati i koja se ne može nadzirati na transparentan način, između ostalog zbog prava intelektualnog vlasništva, ne smije biti certificirana.

Privatni sektor trebao bi se organizirati ili ujediniti kako bi stvorio oznake kvalitete za sustave UI-ja s ciljem stvaranja pozitivnog kruga za te proizvode kako bi tijela koje rade na sprečavanju, otkrivanju ili istraživanju kaznenih djela lakše utvrdila koji sustavi UI-ja zadovoljavaju njihove potrebe.

10. Sva ljudska prava moraju biti zaštićena kada se koriste sustavi UI-ja za sprečavanje, otkrivanje ili istraživanje kaznenih djela. Države te regionalna i međunarodna tijela moraju osigurati izricanje učinkovitih, proporcionalnih i odvrćajućih sankcija kada se takva prava prekrše.

Zakoni ili norme jednake pravne snage moraju izričito predvidjeti da, kada je uzrok povrede ljudskih prava tehnološki kvar sustava UI-ja, tvrtka koja je stvorila sustav snosi odgovornost za grešku ili nemar ili temeljem stroge odgovornosti za neispravne proizvode. Također, moraju osigurati provođenje istraga kako bi se utvrdio uzrok povrede.

11. Sve sadašnje rezolucije primjenjuju se i na sprečavanje, otkrivanje, istraživanje i sankcioniranje upravnih prekršaja od strane nadležnih tijela.

Rezolucije specifične za „prognostički policijski rad“

12. Države te regionalna i međunarodna tijela za ljudska prava moraju osigurati da upotreba sustava UI-ja u sprečavanju i otkrivanju kaznenih djela ne dovede do masovnog nadzora, što bi rezultiralo nerazmjernim ograničavanjem individualnih sloboda (slobode kretanja, izražavanja, okupljanja, udruživanja i vjeroispovijesti).

Konkretno, države i lokalne vlasti moraju zabraniti upotrebu sustava UI-ja za daljinsku identifikaciju osoba na temelju njihovih biometrijskih podataka u javno dostupnim prostorima, kao i bilo koju drugu upotrebu tih sustava koja omogućuje masovni nadzor.

Države se pozivaju da budu transparentnije u vezi s korištenjem sustava automatskog prepoznavanja registarskih tablica u javno dostupnom prostoru. Kada ti sustavi ne uključuju samo fotografiranje registarske tablice već i slikanje bilo koje osobe u vozilu, ta mogućnost mora biti izričito dopuštena zakonom. Primjena tehnologije prepoznavanja lica na podatke prikupljene putem tih slika mora biti zabranjena za potrebe prognostičkog policijskog rada. Ona se može primijeniti samo u okviru specifične istrage ako za to postoji pravni okvir.

13. Države moraju utvrditi ili osigurati da neovisna istraživačka tijela utvrde pomaže li upotreba sustava UI-ja u sprečavanju kaznenih djela i smanjenju broja počinjenih kaznenih djela te, ako pomaže, u kojem postotku.
14. Države moraju osigurati da zbog financijskog troška sustava UI-ja i njihova održavanja ne uskraćuju sredstva javnim službama za prevenciju kriminala koje se bave uzrocima kriminala (psihološka podrška, socijalna podrška, osposobljavanje i pomoć u zapošljavanju).
15. Zakoni i norme jednake pravne snage moraju strogo zabraniti upotrebu podataka kao dokaza optužbe u kaznenim postupcima ako su ti podaci prikupljeni putem sustava UI-ja u vezi s prevencijom kriminala, odnosno u situacijama kada nije postojala konkretna sumnja da je počinjeno kazneno djelo i stoga su podaci prikupljeni izvan zakonskog okvira koji uređuje kaznene istrage. Ako se podaci prikupljeni putem sustava UI-ja u kontekstu prevencije kriminala koriste kao osnova za pokretanje istrage („početne informacije“) ili kao početna informacija za kaznenu istragu, nadležno pravosudno tijelo mora biti obaviješteno o tome. Ti podaci moraju biti označeni kao takvi, a upotreba sustava UI-ja mora biti dokumentirana u spisu predmeta.

Rezolucije specifične za „prognostičko pravosuđe“

16. Zakoni i norme jednake pravne snage moraju strogo zabraniti upotrebu sustava UI-ja u svrhu aktuarskog pravosuđa pri izricanju kazni. Kažnjavanje ili pooštavanje kazne na temelju vjerojatnosti da će netko u budućnosti počiniti kazneno djelo znači primjenu kazne dijelom na osnovi kaznenog djela koje se nije dogodilo. To je protivno ljudskom dostojanstvu, osobnoj slobodi i temeljnim načelima kaznenog pravosuđa. Korištenje alata UI-ja za procjenu rizika mora biti zabranjeno kada su u pitanju ozbiljne mjere opreza, poput istražnog zatvora. Kada države dopuštaju korištenje takvih alata za blaže mjere, zakon ih mora izričito odobriti, uz dovoljna procesna jamstva. Međutim, vjerojatnosti koje daje UI ne mogu biti jedini temelj za donošenje odluke.
17. Države koje žele koristiti UI za pomoć tužiteljicama i/ili sutkinjama u kvantitativnim pravnim analizama prije donošenja odluka u kaznenim predmetima moraju ograničiti upotrebu te tehnologije na manje prekršaje koji predstavljaju velik broj slučajeva.
18. Prije nego što odluče koristiti UI za olakšavanje upravljanja velikim brojem slučajeva koji uključuju manje prekršaje, države moraju procijeniti bi li, s obzirom na načelo *ultima ratio*, bilo primjereno dekriminalizirati ponašanje koje generira takve slučajeve.

19. Zakoni i norme jednake pravne snage moraju zabraniti upotrebu kvantitativne pravne analize za pomoć sutkinjama prilikom donošenja odluka o krivnji.
20. Zakoni i norme jednake pravne snage moraju zabraniti upotrebu kvantitativne pravne analize za pomoć sutkinjama pri izricanju kazni. Odluku o kažnjavanju osobe i vrsti kazne moraju donositi ljudi. U suprotnom, suđenje se može dehumanizirati, a ljudsko dostojanstvo može biti ugroženo.
21. Zakoni i norme jednake pravne snage moraju zabraniti upotrebu kvantitativne pravne analize za pomoć sutkinjama prilikom donošenja odluka u kaznenim stvarima koje se donose prije presude, a koje uključuju prisilne mjere.
22. Države moraju osigurati da odluke donesene uz pomoć kvantitativne pravne analize ne krše pravo na pristup ljudskoj sutkinji.
23. Zakoni i norme jednake pravne snage moraju zabraniti pomoć kvantitativne pravne analize osim ako se odluka može pobijati žalbom od strane osobe na koju se odnosi. Odluka o žalbi ne smije biti zasnovana isključivo na kvantitativnoj pravnoj analizi.

Rezolucije specifične za dokaze prikupljene i/ili proizvedene od strane UI-ja

24. Zakoni i norme jednake pravne snage o prikupljanju podataka za analizu sustavom UI-ja moraju zahtijevati da, prije nego što od osobe zatraži pristupni kod njezina softvera ili hardvera iz kojeg se mogu izvući podaci, nadležno tijelo mora obavijestiti tu osobu o njezinu pravu da ne inkriminira samu sebe.
25. Zakoni i norme jednake pravne snage o analizi kriminala moraju odrediti da dijagrami analize kriminala proizvedeni od strane sustava UI-ja nemaju dokaznu vrijednost, ali mogu služiti kao vodič za vođenje istrage.
26. Zakoni i norme jednake pravne snage o uporabi sustava UI-ja za prikupljanje dokaza ili proizvodnju podataka za svrhe kaznenog pravosuđa moraju jasno upozoriti na to da su rezultati sustava UI-ja samo vjerojatnosti. Moraju zahtijevati da sve odluke utemeljene na vjerojatnosti navedu ne samo procjenu vjerojatnosti koju je dao sustav UI-ja već i stopu pogreške tog sustava, kako je izračunala certificirana organizacija koja ga je procijenila.
27. Države i pravosudna tijela moraju osigurati da korištenje UI-ja u izračunu vjerojatnosti ne snižava postojeći standard dokazivanja u kaznenim postupcima.
28. Zakoni i norme jednake pravne snage o uporabi sustava UI-ja za prikupljanje dokaza ili proizvodnju podataka za svrhe kaznenog pravosuđa moraju zabraniti upotrebu, kao dokaza, vjerojatnosti koje je izračunao sustav UI-ja koji nije u potpunosti objašnjiv (crna kutija UI-ja).
29. Zakoni i norme jednake pravne snage o uporabi sustava UI-ja za prikupljanje dokaza ili proizvodnju podataka za svrhe kaznenog pravosuđa moraju zahtijevati, u skladu s pravom na kontradiktorno suđenje, da, ako se koriste podaci prikupljeni ili proizvedeni od sustava UI-ja, sve strane budu obaviještene o tome. Podaci moraju biti označeni kao takvi, a uporaba sustava UI-ja mora biti dokumentirana u spisu predmeta.

30. Zakoni i norme jednake pravne snage moraju postaviti načelo da strana koja predlaže vjerojatnost pred sudom mora sustavno uključiti potpune informacije o tome kako sustav UI-ja funkcionira i koje podatke koristi.
31. Zakoni i norme jednake pravne snage o uporabi sustava UI-ja za prikupljanje dokaza ili proizvodnju podataka za svrhe kaznenog pravosuđa moraju, u skladu s pravima obrane, omogućiti da svaka osoba optužena na osnovi vjerojatnosti koja se koristi kao dokaz može dobiti izvorni kod sustava UI-ja i podatke za obuku kako bi ih stručnjak mogao analizirati. Poslovna tajna ne smije ugroziti prava obrane.
32. Zbog visokih troškova dobivanja stručne analize sustava UI-ja države moraju osigurati da svatko tko je optužen na osnovi vjerojatnosti izračunate od strane sustava UI-ja ima pristup ne samo učinkovitoj pravnoj pomoći već i financijskoj pomoći za takvu specifičnu ekspertizu.

IV. SEKCIJA: MEĐUNARODNE PERSPEKTIVE O UI-ju: IZAZOVI ZA PRAVOSUDNU SURADNJU I MEĐUNARODNO HUMANITARNO/ KAZNENO PRAVO****

Rezolucija je usvojena na međunarodnom kolokviju „Međunarodne perspektive o UI-ju: izazovi za pravosudnu suradnju i međunarodno humanitarno/kazneno pravo“, Opatija, Hrvatska, 6.–8. prosinca 2023.

Preambula

Uzimajući u obzir da je umjetna inteligencija (UI) sastavni dio procesa donošenja odluka autonomnih oružanih sustava (AOS, engl. AWS), a takva su oružja već postigla značajnu razinu razvoja u ovom trećem desetljeću 21. stoljeća te ih mnoge države i nedržavni dionici već široko koriste;

Uzimajući u obzir da su višestruke odluke o nadzoru i ciljanju, koje tradicionalno donose ili ih oblikuju ljudi, počele biti automatizirane upotrebom AOS-a, uzrokujući značajna pravna, moralna i etička razmatranja;

Ističući sve veću zabrinutost zbog štetnosti zlonamjernih ili nepažljivih upotreba AOS-a;

Uvažavajući potrebu za analizom je li pravni odgovor država i međunarodne zajednice na izazove UI-ja dovoljan ili ga je potrebno reformirati i prilagoditi, bilo specifičnim izmjenama bilo stvaranjem novih oblika kriminalizacije;

Uvažavajući potrebu da AOS-i budu dizajnirani i razvijeni u skladu s međunarodnim pravom.

Primjećujući da se AOS-e može definirati kao „[b]ilo koji oružani sustav s autonomijom u svojim ključnim funkcijama – tj. oružani sustav koji može odabrati (tražiti, otkriti, identificirati, pratiti ili odabrati) i napasti (koristiti silu protiv, neutralizirati, oštetiti ili uništiti) ciljeve bez ljudske intervencije”;²

Primjećujući međutim da ta definicija AOS-a nije univerzalno prihvaćena te da različiti međunarodni, regionalni i nacionalni režimi mogu odabrati različita shvaćanja definicije AOS-a i smislene ljudske kontrole pri njihovu korištenju („čovjek u petlji”);

**** Rezoluciju je preveo dr. sc. Marin Bonačić, izvanredni profesor Pravnog fakulteta Sveučilišta u Zagrebu, pri čemu se kao pomoćnim alatom teksta koristio platformom UI-ja ChatGPT, UI Microsoft Copilot i Google Translate. Prijevod je temeljito provjeren, ispravljen, kontekstualiziran i usklađen sa standardnom pravnom terminologijom.

² Međunarodni odbor Crvenoga križa, Stavovi MOCK-a o autonomnim oružanim sustavima (*Views of the ICRC on autonomous weapon systems*), dokument podnesen Konvenciji o određenim konvencionalnim oružjima, sastanku stručnjaka o smrtonosnim autonomnim oružanim sustavima (*Convention on Certain Conventional Weapons Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*), 11. travnja 2016., <https://www.icrc.org/en/document/views-icrc-autonomous-weapon-system>.

Primjećujući da različiti međunarodni, regionalni i nacionalni režimi također mogu odabrati razlikovati AOS-e i smrtonosne autonomne oružane sustave te također AOS-e i automatizirane oružane sustave;

Uzimajući u obzir da upotreba AOS-a može uzrokovati značajnu štetu ljudima i društvima te da takva upotreba izaziva pravne i etičke posljedice povezane s *ius ad bellum* i *ius in bello*, teškim kršenjima ljudskih prava, kršenjima *ius cogens* i temeljnim (engl. *core*) međunarodnim zločinima;

Uzimajući u obzir da upotreba AOS-a može utjecati na javno mnijenje i politiku u korist upotrebe sile, zbog percepcije da upotreba AOS-a smanjuje rizike od uništenja, smrti ili tjelesne ozljede vojnika i drugih uključenih pojedinaca;

Uzimajući u obzir da upotreba AOS-a može negativno utjecati na poštovanje temeljnih načela *ius in bello*, kao što su načela razlikovanja, razmjernosti i opreza;

Uzimajući u obzir da uklanjanjem ili smanjenjem ljudskog elementa u donošenju odluka upotreba AOS-a može pridonijeti povećanju smrtnih slučajeva zbog odsutnosti ljudskih osjećaja, poput straha i suosjećanja, koji mogu igrati ulogu u smanjenju smrtnih ishoda;

Uzimajući u obzir da upotreba AOS-a može uzrokovati značajno uništenje i kolateralnu štetu;

Uzimajući u obzir da se upotrebom AOS-a mogu počiniti temeljni međunarodni zločini te da takva upotreba AOS-a može otvoriti ozbiljna i nova pitanja u vezi s atribucijom kaznene odgovornosti, uključujući i pitanja u svezi sa zapovjednom odgovornosti, ali ne ograničavajući se samo na njih.

Uzimajući u obzir da upotreba AOS-a može otvoriti jurisdikcijska pitanja, jer upotreba AOS-a može biti transteritorijalna;

Utvrđujući da je nužan međunarodni regulatorni pristup AOS-ima i da bi trebao uključivati pravno obvezujuću normu koja zabranjuje dizajn, razvoj, proizvodnju i upotrebu potpuno autonomnih AOS-a (bez smislene ljudske kontrole);

Potičući sve države da otvore službene pregovore unutar Ujedinjenih naroda (pod pokroviteljstvom Konvencije o zabrani ili ograničenju uporabe određenog konvencionalnog oružja i/ili unutar drugih odgovarajućih konvencija, tijela, agencija i institucija UN-a) s ciljem razvijanja propisa ili drugih smjernica koji će se primjenjivati na dizajn, razvoj, proizvodnju i upotrebu AOS-a.

Utvrđujući da to također povećava potrebu za globalnim pristupom AOS-ima.

Svjesni:

*Izvjешća o neformalnom sastanku stručnjaka o smrtonosnim autonomnim oružanim sustavima (LAWS) Ujedinjenih naroda 2016. godine;*³

³ [https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_\(2016\)/ReportLAWS_2016_AdvancedVersion.pdf](https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2016)/ReportLAWS_2016_AdvancedVersion.pdf).

Konvencije o zabrani ili ograničenju uporabe određenog konvencionalnog oružja s pretjeranim traumatskim učinkom ili djelovanjem bez obzira na cilj Ujedinjenih naroda, kako je izmijenjena 21. prosinca 2001.;⁴

Rezolucije Europskog parlamenta o autonomnim oružanim sustavima od 12. rujna 2018.;⁵

Latinskoameričke i karipske konferencije o socijalnom i humanitarnom utjecaju autonomnih oružja, Priopćenje Latinskoameričke i karipske konferencije o socijalnom i humanitarnom utjecaju autonomnih oružja, 23. i 24. veljače 2023.;

Izvrješća Međunarodnog odbora Crvenog križa, Autonomni oružani sustavi: implikacije povećane autonomije u ključnim funkcijama oružja, Ženeva, rujan 2016.;⁶

Izjave na Konvenciji o određenim konvencionalnim oružjima (CCW) na sastanku stručnjaka o smrtonosnim autonomnim oružanim sustavima (LAWS) Međunarodnog odbora Crvenog križa, 13.–17. travnja 2015., Ženeva;⁷

XX. Kongresa AIDP-a, Kazneno pravosuđe i korporativno poslovanje, Rim, 13.–16. studenog 2019.⁸

REZOLUCIJE

- 1) Kako bi se spriječile i smanjile štete povezane s AOS-ima, potrebno je da međunarodni, regionalni, nacionalni zakonodavci i druge nadležne vlasti u potpunosti definiraju AOS-e, kao i da razviju propise koji uređuju dizajn, razvoj, proizvodnju i upotrebu AOS-a.
- 2) Upotreba AOS-a mora prethodno biti regulirana zakonom ili normom jednake obvezujuće snage.⁹
- 3) Države se pozivaju na transparentniju upotrebu AOS-a.
- 4) Države moraju osigurati da odluke vlasti o korištenju AOS-a budu utemeljene na transparentnim kriterijima koji podliježu javnom nadzoru. Države koje

⁴ <https://disarmament.unoda.org/the-convention-on-certain-conventional-weapons/#:~:text=It%20was%20adopted%20on%2010,or%20to%20affect%20civilians%20indiscriminately.>

⁵ https://www.europarl.europa.eu/doceo/document/TA-8-2018-0341_EN.html.

⁶ ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons*, MOCK, Ženeva, rujan 2016., <https://www.icrc.org/en/publication-4283-autonomous-weapons-systems>.

⁷ ICRC (2015) *Statement to the Convention on Certain Conventional Weapons (CCW) Meeting of Experts on Lethal Autonomous Weapons Systems (LAWS)*, 13.–17. travnja 2015., Ženeva, <https://www.icrc.org/en/document/lethal-autonomous-weapons-systems-LAWS>.

⁸ <https://www.penal.org/fr/aidp-xx-international-congress-penal-law-%E2%80%9Ccriminal-justice-and-corporate-business%E2%80%9D-rome-13-16-0>.

⁹ Članak 36. Dodatnog protokola I Ženevskim konvencijama već zahtijeva od država članica da nastave procjenjivati bi li razvoj ili upotreba bilo kakva novog oružja bio zabranjen Dodatnim protokolom ili bilo kojim drugim pravilom međunarodnog prava. Dodatni protokol Ženevskim konvencijama od 12. kolovoza 1949., koji se odnosi na zaštitu žrtava međunarodnih oružanih sukoba (Protokol I), članak 36., <https://www.ohchr.org/en/instruments-mechanisms/instruments/protocol-additional-geneva-conventions-12-august-1949-and>.

- koriste AOS-e moraju odabrati sustave čije je djelovanje potpuno transparentno, objašnjivo i sljedivo (UI iz bijele kutije, engl. *white box AI*).
- 5) Zakon ili druge jednakovrijedne norme koje se odnose na korištenje AOS-a moraju zahtijevati da takvi sustavi imaju visoku razinu tehnološke pouzdanosti. Dovoljno precizna regulativa koja zahtijeva odgovarajuće evaluacije i validacije, neovisne o razvojnom programeru AOS-a, mora u najvećoj mogućoj mjeri ograničiti rizik od pristranosti ili bilo kojeg oblika diskriminacije u strojnom učenju, grešaka u kodiranju i drugih tehnoloških kvarova.
 - 6) Zakoni ili jednakovrijedne norme moraju zahtijevati da se, prije nego što se AOS-i utemeljeni na samoučećim algoritmima mogu koristiti za nadzor, odabir ciljeva ili bilo koju drugu svrhu, moraju razviti, obučiti, testirati i koristiti algoritmi pod ljudskim nadzorom (načelo čovjeka u petlji). Ti zakoni i jednakovrijedne norme moraju zahtijevati ljudsku procjenu prije poduzimanja bilo kakve radnje u svrhu korištenja AOS-a.
 - 7) Države i tijela za provedbu zakona moraju osigurati da njihovo osoblje koje upravlja AOS-ima dobije odgovarajuću obuku o pravilnoj upotrebi relevantnih AOS-a, kao i obuku u svezi s rizikom od pogreške i pristranosti. Moraju osigurati da takvo osoblje ima temeljito znanje o opasnostima koje AOS-i mogu predstavljati za ljudska prava.
 - 8) Zakoni ili jednakovrijedne norme moraju izričito propisati da u slučaju kršenja ljudskih prava kao rezultata bilo kakve neispravnosti AOS-a tvrtka koja je dizajnirala ili proizvela sustav snosi kaznenu odgovornost zbog pogreške ili nehaja,¹⁰ bez isključivanja građanske odgovornosti za neispravne proizvode. Takvi zakoni ili jednakovrijedne norme također moraju osigurati obvezu provođenja istraga kako bi se utvrdio uzrok kršenja.
 - 9) Svaka država koja odluči koristiti AOS-e mora imati odgovarajuće kaznene zakone koji moraju predvidjeti da će osobe koje koriste AOS-e na način koji nije u skladu s relevantnim međunarodnim i nacionalnim pravnim standardima snositi individualnu kaznenu odgovornost pred nadležnim kaznenim jurisdikcijama.¹¹
 - 10) Države te regionalna i međunarodna tijela za ljudska prava moraju osigurati da upotreba AOS-a ne dovede do teških kršenja ljudskih prava, kršenja *ius cogens* ili drugih kršenja relevantnih normi međunarodnog prava i međunarodnog humanitarnog prava.
 - 11) Države moraju osigurati potpunu implementaciju svojih međunarodnopravnih obveza u svoje nacionalne pravne sustave putem zakona, uredbi, izvršnih

¹⁰ XXI. Kongres AIDP-a: Umjetna inteligencija i kazneno pravosuđe, I. sekcija: Tradicionalni instituti kaznenog prava i UI, rezolucije prihvaćene na Kolokviju u Siracusi, 15.–17. rujna 2022.; XXI. Kongres AIDP-a: Umjetna inteligencija i kazneno pravosuđe: II sekcija: Kazneno pravo i kriminalizacija suočena s izazovima UI-ja, rezolucije prihvaćene na Kolokviju u Bukureštu, 14.–16. lipnja 2023.

¹¹ XXI. Kongres AIDP-a: Umjetna inteligencija i kazneno pravosuđe, I. sekcija: Tradicionalni instituti kaznenog prava i UI, rezolucije prihvaćene na Kolokviju u Siracusi, 15.–17. rujna 2022.

naredbi, vojnih kodeksa ili drugih vrsta obvezujućih domaćih pravnih normi. Takvo domaće zakonodavstvo mora predvidjeti različite načine odgovornosti, uključujući i zapovjednu odgovornost, ali ne ograničavajući se samo na nju.

- 12) Države trebaju uspostaviti neovisna istražna tijela koja će kontinuirano proučavati provodi li se upotreba AOS-a na način koji je u skladu sa svim primjenjivim međunarodnim i nacionalnim pravnim normama i propisima.
- 13) Potrebno je identificirati i definirati specifične načine atribucije kaznene odgovornosti osobama (kako fizičkim tako i pravnim) koje su odgovorne za razvoj AOS-a.¹² Takva odgovornost ne bi trebala isključivati osobe, bilo fizičke bilo pravne, koje pridonose uzročnom lancu štete: od dizajnera, programera, proizvođača, prodavatelja, distributera do samih krajnjih korisnika sustava.
- 14) Države bi trebale težiti poticanju suradnje između političkih i akademskih stručnjaka o temi međunarodnog prava te pripadnika vojske, tijela za provedbu zakona ili obavještajnih službi koji su zaduženi za razvoj i/ili raspoređivanje AOS-a. Takva suradnja trebala bi se provoditi s ciljem osiguranja da je upotreba AOS-a uvijek usklađena s relevantnim međunarodnopravnim normama, kao i s primjenjivim domaćim zakonima i propisima.
- 15) Države moraju osigurati eksteritorijalnu primjenu svojih domaćih zakona u slučajevima kada se AOS-i raspoređuju ili na drugi način koriste u inozemstvu od strane država ili entiteta pod kontrolom države.
- 16) Države moraju donijeti odgovarajuće zakone o sukobu zakona i sukobu jurisdikcija koji omogućuju primjenu domaćeg prava u slučajevima kada se AOS-i raspoređuju ili na drugi način koriste u inozemstvu. Takvi zakoni mogu se donijeti na domaćoj, regionalnoj i/ili međunarodnoj razini. Takvi zakoni moraju omogućiti uspostavljanje kaznene odgovornosti (kao i građanske odgovornosti) protiv onih fizičkih ili pravnih osoba pod kontrolom relevantne države koje su odgovorne za zloupotrebu ili neispravan rad AOS-a smještenih izvan teritorija relevantne države.
- 17) Države moraju razviti modele pravosudne suradnje u kaznenim stvarima, posebno kako bi prikupile dokaze u svezi s kaznenim djelima uzrokovanim korištenjem AOS-a. Takvi modeli pravosudne suradnje mogu biti posebno potrebni državama koje još nisu zaključile odgovarajuće bilateralne, regionalne ili međunarodne sporazume o toj temi.
- 18) Države moraju osigurati postojanje mehanizama izručenja kako bi učinkovito procesuirale i kaznile one koji su odgovorni za kaznena djela uzrokovana korištenjem AOS-a.

¹² XXI. Kongres AIDP-a: Umjetna inteligencija i kazneno pravosuđe, I. sekcija: Tradicionalni instituti kaznenog prava i UI, rezolucije prihvaćene na Kolokviju u Siracusi, 15.–17. rujna 2022.