

EVROPSKA KOMISIJA ZA EFIKASNOST PRAVOSUĐA (CEPEJ)

Evropska etička
povelja o korištenju umjetne
inteligencije u pravosuđu



Usvojena na 31. plenarnoj sjednici
CEPEJ-a (Strasbourg, 3-4. decembar 2018)



EVROPSKA KOMISIJA ZA EFIKASNOST PRAVOSUĐA (CEPEJ)

Evropska etička povelja o korištenju umjetne inteligencije u pravosuđu

Usvojena na 31. plenarnoj sjednici
CEPEJ-a (Strasbourg, 3-4. decembar 2018)

French
edition: *Commission
européenne pour l'efficacité de la
justice (CEPEJ) – Charte éthique
européenne d'utilisation de
l'intelligence
artificielle dans les
systèmes judiciaires et leur
environnement*

*Mišljenja iznesena u ovoj
publikaciji su odgovornost autora
i ne predstavljaju nužno službenu
politiku Vijeća Evrope.*

Svi zahtjevi u vezi sa
reprodukcijom ili prevodom
cjelokupnog dokumenta ili
nekog njegovog dijela trebaju
se uputiti Direktoratu za
komunikacije (F-67075
Strasbourg Cedex ili na
adresi publishing@coe.int).
Svu ostalu korespondenciju
u vezi sa ovim dokumentom
treba uputiti Evropskoj
komisiji za efikasnost
pravosuđa (CEPEJ)
cepej@coe.int

Dizajn i
grafička priprema: Odjel
za dokumente i
publikacije (SPDP),
Vijeće Evrope
Fotografije:
Shutterstock

Ova publikacija nije
lektorirana u Jedinici za

lektorisanje SPDP-a radi
ispravke štamparskih i
gramatičkih grešaka.

© Vijeće Evrope, februar
2019. Štampano u Vijeću
Evrope

Sadržaj

- | | | |
|----|---|---|
| 1. | Princip poštovanja temeljnih prava: garantirati usklađenost razvoja i primjene alata i usluga umjetne inteligencije sa temeljnim pravima | 8 |
| 2. | Princip nediskriminacije: posebno spriječiti nastanak ili pojačavanje diskriminacije između pojedinaca ili grupa | |
| 3. | Princip kvaliteta i sigurnosti: prilikom obrade sudskih odluka i podataka koristiti certificirane izvore i modele koji su razvijeni multidisciplinarno i u sigurnom tehnološkom okruženju | |
| 4. | Princip transparentnosti, nepristrasnosti i pravičnosti: voditi računa da metode obrade podataka budu pristupačne i razumljive, predvidjeti vanjsku kontrolu | |
| 5. | Princip „pod kontrolom korisnika“: spriječiti preskriptivni pristup i osigurati da korisnici budu informirani akteri i da imaju mogućnost izbora | |
-
-
-
-
-
-
-
-
-
-
-

Uvod

Uviđajući sve veći značaj umjetne inteligencije¹ (UI) u našem modernom društvu, kao i očekivanih prednosti njene pune primjene u službi efikasnosti i kvaliteta pravosuđa, CEPEJ formalno donosi 5 temeljnih principa pod nazivom „Evropska povelja o korištenju UI u pravosuđu“.

Ova povelja je namijenjena akterima iz javnog i privatnog sektora koji su zaduženi za razvoj i uvođenje alata i usluga umjetne inteligencije koji podrazumijevaju obradu sudskih odluka i podataka (mašinsko učenje ili bilo koje druge metode izvedene iz nauke o podacima).

Ona je također relevantna za donosiocje odluka u javnom sektoru odgovorne za zakonodavni ili regulatorni okvir, razvoj, kontrolu ili korištenje tih alata i usluga.

Primjena tih alata i usluga u pravosuđu ima za cilj unapređenje efikasnosti i kvaliteta pravosuđa i treba je podsticati, ali ona mora biti odgovorna i u skladu sa temeljnim pravima građana predviđenim Evropskom konvencijom o ljudskim pravima i Konvencijom o zaštiti ličnih podataka, kao i sa ostalim temeljnim principima definiranim u nastavku koji bi trebali biti okosnica donošenja javnih politika u pravosuđu u ovoj oblasti.

Po mišljenju programera umjetne inteligencije obrada sudskih odluka pomoću umjetne inteligencije moguća je u građanskim, privrednim i upravnim postupcima gdje ona može pomoći u unapređenju predvidljivosti primjene zakona i dosljednosti sudskih odluka, pod uslovom da se poštuju principi definirani u daljem tekstu. Razmatranju njene primjene u krivičnim stvarima treba prići vrlo obazrivo kako bi se spriječila diskriminacija na osnovu osjetljivih podataka uz poštovanje garancija pravičnosti postupka.

Bez obzira na to da li se koristi kao pomoć u pružanju pravnih savjeta, izradi pismenih otpravaka ili u postupku donošenja odluka, odnosno u svrhu pružanja savjeta korisniku, ta obrada mora biti transparentna, nepristrasna i pravična te certificirana u postupku vanjske i nezavisne stručne procjene.

Primjena Povelje

Akteri iz javnog i privatnog sektora bi principe iz Povelje trebali redovno primjenjivati, pratiti i ocjenjivati radi stalnog unapređenja praksi.

U tom smislu, poželjno bi bilo da ovi akteri provode redovnu analizu primjene principa utvrđenih Poveljom praćenu, tamo gdje je to potrebno, obrazloženjem razloga za njihovo neprimjenjivanje ili djelimičnu primjenu te planovima za uvođenje neophodnih mjera.

Nezavisni organi pomenuti u Povelji bi mogli biti zaduženi za periodičnu procjenu stepena poštovanja principa iz Povelje od strane svih aktera, kao i za davanje prijedloga za unapređenje u cilju njenog prilagođavanja razvoju tehnologija i njihovoj primjeni.

Pet principa Etičke povelje o korištenju umjetne inteligencije u pravosuđu

1

Princip poštovanja temeljnih prava: garantirati usklađenost razvoja i primjene alata i usluga umjetne inteligencije sa temeljnim pravima.

2

Princip nediskriminacije: posebno spriječiti nastanak ili pojačavanje diskriminacije između pojedinaca ili grupa.

3

Princip kvaliteta i sigurnosti: prilikom obrade sudskih odluka i podataka koristiti certificirane izvore i modele koji su razvijeni multidisciplinarno i u sigurnom tehnološkom okruženju.

4

Princip transparentnosti, nepristrasnosti i pravičnosti: voditi računa da metode obrade podataka budu pristupačne i razumljive, predvidjeti vanjsku kontrolu.

Princip „pod kontrolom korisnika“: spriječiti preskriptivni pristup i osigurati da korisnici budu informirani akteri i da imaju mogućnost izbora.

1

Princip poštovanja temeljnih prava: garantirati usklađenost razvoja i primjene alata i usluga umjetne inteligencije sa temeljnim pravima

■ Obrada sudskih odluka i podataka mora imati jasnu svrhu i biti u skladu sa temeljnim pravima predviđenim Evropskom konvencijom o ljudskim pravima (EKLJP) i Konvencijom o zaštiti ličnih podataka (Konvencija o zaštiti pojedinaca u vezi sa automatskom obradom ličnih podataka, ETS br. 108 izmijenjena Protokolom CETS br. 223).

■ U slučaju kada se alati vještačke inteligencije koriste za rješavanje sporova ili kao pomoć u donošenju sudskih odluka ili za davanje smjernica građanima mora se voditi računa da oni ne ugrožavaju garantiranje prava na pristup sudiji ili prava na pravično suđenje (jednakost u postupku i poštovanje adversarnog postupka).

■ Kod korištenja tih alata također treba poštovati principe vladavine prava i nezavisnosti sudija u odlučivanju.

■ Prednost stoga treba dati pristupu izradi rješenja zasnovanom na etičkim principima (*ethical-by-design*)² ili poštovanju ljudskih prava (*human- rights-by-design*). To podrazumijeva punu integraciju pravila kojima se zabranjuje direktno ili indirektno kršenje temeljnih vrijednosti zaštićenih konvencijama već u fazi razvoja ili učenja.

2. Etički izbori se prave na nivou dizajnera programa i ne prepuštaju se korisniku.

2

Princip nediskriminacije: posebno spriječiti nastanak ili pojačavanje diskriminacije između pojedinaca ili grupa

■ S obzirom na to da se ovim metodama obrade može razotkriti postojeća diskriminacija kroz grupisanje ili klasifikaciju podataka o pojedincima ili grupama, akteri iz javnog i privatnog sektora moraju osigurati da te metode ne reproduciraju ili pogoršaju tu diskriminaciju te da ne dovedu do determinističkih analiza ili primjene.

■ Posebnu pažnju treba obratiti u fazi razvoja i uvođenja, naročito kada je obrada direktno ili indirektno zasnovana na „osjetljivim podacima“ u koje mogu spadati rasna ili etnička pripadnost, socio-ekonomski položaj, politički stavovi, vjerska ili filozofska uvjerenja, članstvo u sindikatima, genetički podaci, biometrijski podaci, podaci o zdravstvenom stanju ili podaci o seksualnom životu ili seksualnoj orijentaciji. U slučaju otkrivanja diskriminacije neophodno je razmotriti preduzimanje korektivnih mjera u cilju ograničavanja, ili po mogućnosti, neutralizacije tih rizika, kao i u cilju podizanja svijesti aktera.

■ Međutim, korištenje mašinskog učenja i multidisciplinarnih naučnih analiza za borbu protiv diskriminacije treba podsticati.

3

Princip kvaliteta i sigurnosti: prilikom obrade sudskih odluka i podataka koristiti certificirane izvore i modele koji su razvijeni multidisciplinarno i u sigurnom tehnološkom okruženju

■ Dizajneri modela mašinskog učenja bi trebali biti u prilici da sarađuju sa profesionalcima iz pravosudnog sistema (sudije, tužiocima, advokati itd.) i naučnicima/predavačima iz oblasti prava i društvenih nauka (npr. ekonomistima, sociolozima i filozofima).

■ Jedna od organizacionih metoda za iskorištavanje prednosti ovog multidisciplinarnog pristupa je osnivanje mješovitih projektnih timova koji će u kratkim ciklusima izrade rješenja raditi na razvoju funkcionalnih modela.

■ Ovi projektni timovi bi trebali stalno izvještavati o etičkim mehanizmima i unapređivati ih na osnovu povratnih informacija.

■ Podaci zasnovani na sudskim presudama koji se unose u softver koji koristi algoritam mašinskog učenja bi trebali dolaziti iz certificiranih izvora i ne bi se smjeli modificirati prije korištenja u mehanizmu za učenje. Cijeli proces bi, dakle, trebao biti sljediv kako bi se onemogućila bilo kakva izmjena koja može uticati na sadržaj ili značenje odluke koja se obrađuje.

■ Kreirani modeli i algoritmi se također moraju čuvati i primjenjivati u sigurnom okruženju radi garantiranja integriteta sistema.

4

Princip transparentnosti, nepristrasnosti i pravičnosti: voditi računa da metode obrade podataka budu pristupačne i razumljive, predvidjeti vanjsku kontrolu

— Kod primjene alata koji mogu imati pravne posljedice ili mogu značajno uticati na život ljudi neophodno je naći sredinu³ između prava intelektualnog vlasništva u vezi sa određenim metodama obrade i potrebom za transparentnošću (pristup procesu izrade rješenja), nepristrasnosti (nepostojanje subjektivnosti)⁴, pravičnost i intelektualni integritet (stavljanje interesa pravosuđa na prvo mjesto). Mora biti jasno precizirano da se te mjere odnose na cjelokupan lanac izrade i funkcioniranja rješenja pošto proces odabira i kvalitet i organizacija podataka imaju direktan uticaj na fazu učenja.

— Prva opcija je potpuna tehnička transparentnost (na primjer kod i dokumentacija iz otvorenog izvora), što je u nekim slučajevima ograničeno zaštitom poslovne tajne. Druga opcija bi bila da se sistem objasni jasnim i poznatim jezikom (opis načina dobijanja rezultata) kroz, na primjer, pružanje informacija o prirodi usluga koje se nude, razvijenim alatima, funkcioniranju i riziku od grešaka. Mogli bi se imenovati nezavisni organi ili eksperti koji bi bili zaduženi za certifikaciju i kontrolu metoda obrade ili savjetovanje. Javni organi bi mogli odobravati certifikaciju koja bi podlijeerala redovnim provjerama.

3. U ovom smislu interesantan je prijedlog dat na strani 38 studije MSI-NET-a Vijeća Evrope pod nazivom „Algoritmi i ljudska prava“: Stavljanje na raspolaganje kompletnih algoritama ili osnovnog softverskog koda javnosti je malo izvodljivo rješenje u ovom kontekstu pošto su algoritmi za privatne kompanije glavni vlasnički softver koji je zaštićen. Međutim, možda bi se moglo zahtijevati da određeni se određeni osnovni podskupovi informacija o algoritmima stave na raspolaganje javnosti, na primjer varijable koje se koriste, ciljevi za koje su algoritmi optimizirani, podaci o treniranju i prosječne vrijednosti i standardna odstupanja produciranih rezultata, ili količina i vrsta podataka koje algoritam obrađuje.“ Ili čak prijedlozi dati na strani 117 izvještaja pod nazivom „UI u službi čovječanstva“ koji je po nalogu premijera Francuske Republike pripremio g. Cédric Villani, poslanik u Narodnoj skupštini Francuske: „Kontrolori se mogu zadovoljiti jednostavnom provjerom pravičnosti sistema (uraditi samo ono što se od njih traži), na primjer dostavljanjem različitih lažnih ulaznih podataka, ili kreiranjem velikog broja profila korisnika sistema u skladu sa preciznim smjernicama“. Pored toga, tu su i navodi iz izvještaja Doma lordova pod nazivom „UI u UK: spremni, voljni i sposobni?“, paragraf 92, 96-99.

4. U tom smislu vrijedi napomenuti rješenja za garantiranje neutralnosti algoritama pomenuta u navedenom izvještaju Doma lordova (paragraf 114, 115, 116, 119, 120): raznovrsniji skupovi podataka, veća raznovrsnost i multidisciplinarni pristupi, veća kontrola aspekata kao što su obrada podataka i način projektiranja mašine.

5

Princip „pod kontrolom korisnika“: spriječiti preskriptivni pristup i osigurati da korisnici budu informirani akteri i da imaju mogućnost izbora

■ Primjena alata i usluga umjetne inteligencije mora povećati, a ne ograničiti autonomiju korisnika.

■ Profesionalci iz pravosudnog sistema bi trebali imati mogućnost da, u bilo kom trenutku, revidiraju sudske odluke i podatke koji se koriste za dobijanje rezultata, kao i slobodu da ih ne koriste u svjetlu specifičnih karakteristika datog predmeta.

■ Korisnik mora jasnim i razumljivim jezikom biti informiran od tome da li su rješenja koja nude alati umjetne inteligencije obavezujuća, o različitim opcijama koje imaju na raspolaganju, kao i o tome da imaju pravo na pravni savjet i pravo pristupa sudu. Korisnici također moraju biti jasno informirani o eventualnoj prethodnoj obradi predmeta pomoću umjetne inteligencije prije ili tokom sudskog postupka te moraju imati pravo na prigovor i zahtjev da predmet bude saslušan direktno pred sudom u skladu sa članom 6 EKLJP-a.

■ Načelno govoreći, kod uvođenja informacionog sistema zasnovanog na umjetnoj inteligenciji neophodno je organizirati programe kompjuterske pismenosti za korisnike i debate u kojima će učestvovati profesionalci iz pravosudnog sistema.