Zaštita osobnih podataka i umjetna inteligencija: često postavljana pitanja

Zašto je potrebna regulacija umjetne inteligencije?

Uredba (EU) 2024/1689 Europskog parlamenta i Vijeća od 13. lipnja 2024. o utvrđivanju usklađenih pravila o umjetnoj inteligenciji i o izmjeni uredaba (EZ) br. 300/2008, (EU) br. 167/2013, (EU) br. 168/2013, (EU) 2018/858, (EU) 2018/1139 i (EU) 2019/2144 te direktiva 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828 (dalje u tekstu: Uredba o umjetnoj inteligenciji) u uvodnim i završnim odredbama predviđena postupna primjena pojedinih članaka. Sukladno članku 113. Uredbe o umjetnoj inteligenciji, poglavlja I. i II. počela su se primjenjivati od 2. veljače 2025. godine.

Kako je navedeno u članku 1. stavku 1. Uredbe o umjetnoj inteligenciji, predmetnom se nastoji poboljšati funkcioniranje unutarnjeg tržišta i poduprijeti inovacije, uz istodobno promicanje primjene antropocentrične i pouzdane umjetne inteligencije te osiguravanje zdravlja, sigurnosti i visoke razine zaštite temeljnih prava sadržane u Povelji Europske unije o temeljnim pravima (dalje u tekstu „Povelja"), među ostalim i prava na privatnost i zaštitu osobnih podataka (članci 7. i 8. Povelje).

Široka primjena sustava umjetne inteligencije nosi velik potencijal za ostvarenje društvenih koristi, gospodarskog rasta i jačanje inovacijskog kapaciteta Europske unije te njezine globalne konkurentnosti. Međutim, određene karakteristike pojedinih sustava umetne inteligencije mogu generirati nove rizike, osobito one povezane sa sigurnošću korisnika, uključujući njihovu fizičku sigurnost, ali i zaštitu temeljnih prava. Neki moćni i široko primjenjivani modeli umjetne inteligencije mogu čak predstavljati sistemski rizik.

To stvara pravnu nesigurnost i može dovesti do usporenog prihvaćanja tehnologija umjetne inteligencije od strane javnog i privatnog sektora te građana, prvenstveno zbog manjka povjerenja. Također, različiti regulatorni pristupi na nacionalnoj razini mogli bi uzrokovati fragmentaciju jedinstvenog tržišta.

Kao odgovor na te izazove, bilo je nužno zakonodavno djelovati kako bi se osiguralo pravilno funkcioniranje unutarnjeg tržišta umjetne inteligencije, na način koji uravnotežuje potencijalne koristi i rizike.

Kad počinje primjena Uredbe o umjetnoj inteligenciji?

Poglavlja I. i II. počela su se primjenjivati od **2. veljače 2025. godine**.

Poglavlje I. obuhvaća članke od 1. do 4., kojima se uređuje:

- predmet Uredbe o umjetnoj inteligenciji, odnosno svrha,
- područje primjene, odnosno subjekti koji su obvezni ispuniti zahtjeve Uredbe,
- definicije,
- **pismenost u području umjetne inteligencije.**

Osim obveze utvrđivanja primjenjivosti Uredbe o umjetnoj inteligenciji na poslovne aktivnosti, nužno je poduzeti mjere za osiguranje informiranosti i edukacije o regulatornim zahtjevima u području umjetne inteligencije.

Prilikom provedbe aktivnosti edukacije i podizanja svijesti, potrebno je uzeti u obzir tehničku stručnost, prethodno iskustvo, obrazovnu razinu i stupanj osposobljenosti sudionika, kao i specifični kontekst primjene sustava umjetne inteligencije.

*Saznajte više na poveznicama:

- Često postavljana pitanja- Europska komisija
- Repozitorij edukativnih materijala o umjetnoj inteligenciji

Poglavlje II. definira zabranjene prakse u području umjetne inteligencije koje obuhvaćaju između ostaloga:

- Društveno bodovanje
- Biometrijska kategorizacija
- Korištenje manipulativnih i zavaravajućih tehnika
- Iskorištavanje slabosti određene osobe ili skupina
- Procjena počinjenja kaznenog djela
- Kreiranje/proširenje baze podataka za prepoznavanje lice
- Prepoznavanje emocija na radnom mjestu i u obrazovanju

Također, isto određuje stroge uvjete pod kojima se mogu koristiti sustavi za daljinsku biometrijsku identifikaciju u stvarnom vremenu na javnim mjestima.

*Saznajte više na poveznici:

- **Smjernice Europske komisije o zabranjenim praksama u području umjetne inteligencije**

Druga faza počinje **2. kolovoza 2025**. Tada na snagu stupaju dodatni dijelovi Uredbe, godinu dana prije pune primjene. Posebno, od tog datuma počinju se primjenjivati obveze i mehanizmi vezani uz:

- UI modele Opće namjene: Poglavlje V. Uredbe (Obveze pružatelja modela umjetne inteligencije opće namjenje), čime pružatelji tzv. generativnih ili općih modela umjetne inteligencije (npr. velikih jezičnih modela) moraju početi ispunjavati predviđene obveze.
- Strukturu nadzora i provedbe: Poglavlje VII. (Upravljanje na razini Unije i država članica). To znači da do tog datuma trebaju biti uspostavljena nadzorna tijela (nacionalna i europska) predviđena Uredbom, kao i potrebni odbori i koordinacijski mehanizmi.

- Notifikacijska tijela: Poglavlje III., Odjeljak 4. Uredbe, koji se odnosi na notifikacijska tijela i prijavljena tijela za ocjenu sukladnosti.
- Sankcije i povjerljivost: Poglavlje XII. (kazne), kao i članak 78. Uredbe koji propisuje obveze povjerljivosti za Komisiju, nadzorna tijela i prijavljena tijela u postupanju s povjerljivim podacima.
- ***Napomena:*** *Članak 101. Uredbe, koji propisuje novčane kazne za dobavljače sustava umjetne inteligencije, izuzet je iz ove faze i neće se primjenjivati prije potpune primjene Uredbe.*

Ukratko, od 2. kolovoza 2025. primjenjuju se:

- **poglavlja III. odjeljka 4.** pod nazivom „Tijela koja provode prijavljivanje i prijavljena tijela“ (članci od 28. do 39.)
- **poglavlja V.** „Modeli umjetne inteligencije opće namjene“ (članci od 51. do 56.)
- **poglavlje VII.** „Upravljanje“ (članci od 64. do 70.)
- **članak 78.** „Povjerljivost“
- **poglavlje XII.** „Sankcije“ (članci od 99. do 100)

2. kolovoza 2026. označava **početak opće primjene** Uredbe umjetnoj inteligenciji. Od tog datuma **sve preostale odredbe** postaju primjenjive, osim onih za koje je izričito predviđena odgoda. Ključne implikacije ovog datuma su:

- **Visokorizični UI sustavi:** dobavljači **visokorizičnih UI sustava** (kako su definirani u Poglavlju III.) moraju osigurati usklađenost svojih sustava sa svim propisanim zahtjevima o sigurnosti, upravljanju podacima, transparentnosti, nadzoru i dr. Svi novi visokorizični UI sustavi stavljeni na tržište od tog datuma nadalje podliježu potpunoj provjeri sukladnosti prema Uredbi o umjetnoj inteligenciji.
- **Transparentnost UI i označavanje sadržaja:** radi se o obvezama iz **Poglavlja IV.** za određene UI sustave. To uključuje zahtjev da UI sustavi dizajnirani za interakciju s ljudima otkriju svoj UI identitet te da **generirani ili manipulirani sadržaj** bude označen kao takav. Primjerice, pružatelji generativnih alata umjetne inteligencije koji stvaraju sintetički audio, video, slike ili tekst moraju osigurati da je izlazni sadržaj **automatski označen** i prepoznatljiv kao umjetno generiran. Slično tome, svaki *“deepfake”* (UI generiran ili izmijenjen slikovni, audio ili video sadržaj koji može zavarati o identitetu ili stvarnosti) ***subjekt koji uvodi sustav*** mora popratiti jasnim upozorenjem da je sadržaj umjetno generiran ili izmijenjen. Ove obveze imaju određene iznimke (npr. za sadržaj u umjetničke ili satiričke svrhe, ili za uporabu od strane tijela za provedbu zakona) kako je navedeno u članku 50. Uredbe).

Konačno, neke specifične obveze **odgađaju se do 2. kolovoza 2027.** Najvažnije, Uredba o umjetnoj inteligenciji propisuje da će se **članak 6. stavak 1.** i povezane obveze početi primjenjivati tek od tog datuma. Članak 6. stavak 1. definira kriterije za klasifikaciju UI sustava kao visokorizičnih u slučajevima kada su UI sustavi sigurnosne komponente proizvoda obuhvaćenih određenim zakonima (navedenim u Prilogu I. Uredbe) ili sami ti proizvodi. Ova odgoda do 2027. daje dodatno vrijeme proizvođačima i pružateljima takvih UI sustava (npr. u sektoru automobila,

medicinskih uređaja, strojeva i drugih područja navedenih u Prilogu I) da prilagode postojeće proizvode i sustave novim zahtjevima Uredbe o umjetnoj inteligenciji.

Europska komisija će nakon savjetovanja s Europskim vijećem za umjetnu inteligenciju, a najkasnije 2. veljače 2026., izdati smjernice u kojima se utvrđuje praktična provedba ovog članka u skladu s člankom 96. zajedno sa sveobuhvatnim popisom praktičnih slučajeva upotrebe UI sustava koji su visokorizični i koji nisu visokorizični.

Što se smatra sustavom umjetne inteligencije?

Sustav umjetne inteligencije (UI sustav) znači strojni sustav dizajniran za rad s promjenjivim razinama autonomije i koji nakon uvođenja može pokazati prilagodljivost te koji, za eksplicitne ili implicitne ciljeve, iz ulaznih vrijednosti koje prima, zaključuje kako generirati izlazne vrijednosti kao što su predviđanja, sadržaj, preporuke ili odluke koji mogu utjecati na fizička ili virtualna okruženja.

Ta definicija obuhvaća sedam glavnih elemenata: (1) sustav temeljen na strojevima; (2) koji je dizajniran da djeluje s različitim razinama autonomije; (3) koji može pokazivati prilagodljivost nakon implementacije; (4) i koji, za eksplicitne ili implicitne ciljeve; (5) inferira, na temelju ulaza koji prima, kako generirati izlaze (6) kao što su predikcije, sadržaj, preporuke ili odluke (7) koje mogu utjecati na fizička ili virtualna okruženja.

Definicija sustava umjetne inteligencije usvaja perspektivu temeljenu na životnom ciklusu koja obuhvaća dvije glavne faze: fazu prije implementacije ili 'izgradnje' sustava i fazu nakon implementacije ili 'upotrebe' sustava. Sedam elemenata navedenih u toj definiciji ne mora biti prisutno kontinuirano tijekom obje faze tog životnog ciklusa. Umjesto toga, definicija prepoznaje da se određeni elementi mogu pojaviti u jednoj fazi, ali možda neće postojati u obje faze. Ovaj pristup definiranju sustava umjetne inteligencije odražava složenost i raznolikost ovih sustava, osiguravajući da definicija bude usklađena s ciljevima Uredbe o umjetnoj inteligenciji prilagođavajući se širokom spektru sustava umjetne inteligencije.

*Saznajte više na poveznici:

- Smjernice Europske komisije o definiciji sustava umjetne inteligencije

Osnovna terminologija

„dobavljač“ (*eng. provider*) znači fizička ili pravna osoba, tijelo javne vlasti, javna agencija ili drugo javno tijelo koje razvija UI sustav ili UI model opće namjene ili koji ima razvijen UI sustav ili UI model opće namjene i stavlja ga na tržište ili stavlja UI sustav u upotrebu pod vlastitim imenom ili žigom, uz plaćanje ili besplatno

„subjekt koji uvodi sustav umjetne inteligencije“ (*eng. deployer*) znači fizička ili pravna osoba, tijelo javne vlasti, javna agencija ili drugo javno tijelo koje upotrebljava UI sustav u okviru svoje nadležnosti, osim ako se UI sustav upotrebljava u osobnoj neprofesionalnoj djelatnosti;

„rizik” znači kombinacija vjerojatnosti nastanka štete i težine te štete;

“ovlašteni zastupnik” znači fizička ili pravna osoba koja se nalazi u Uniji ili ima poslovni nastan u Uniji, koju je dobavljač UI sustava ili UI modela opće namjene pisanim putem ovlastio da u njegovo ime izvršava i provodi obveze i postupke utvrđene u ovoj Uredbi i koja je takvo ovlaštenje prihvatila;

“uvoznik” znači fizička ili pravna osoba koja se nalazi u Uniji ili ima poslovni nastan u Uniji i koja stavlja na tržište UI sustav s imenom ili žigom fizičke ili pravne osobe s poslovnim nastanom u trećoj zemlji;

”distributer” znači fizička ili pravna osoba u opskrbnom lancu koja nije dobavljač ni uvoznik i koja stavlja UI sustav na tržište Unije;

„operater” znači dobavljač, proizvođač proizvoda, subjekt koji uvodi sustav, ovlašteni zastupnik, uvoznik ili distributer;

“stavljanje na tržište” znači prvo stavljanje UI sustava ili UI modela opće namjene na raspolaganje na tržištu Unije;

„tijelo koje provodi prijavljivanje” znači nacionalno tijelo odgovorno za utvrđivanje i provedbu postupaka potrebnih za ocjenjivanje, imenovanje, obavješćivanje i praćenje tijelâ za ocjenjivanje sukladnosti;

„ocjenjivanje sukladnosti” znači postupak kojim se dokazuje jesu li ispunjeni zahtjevi utvrđeni u poglavlju III. odjeljku 2. koji se odnose na visokorizični UI sustav;

„tijelo za ocjenjivanje sukladnosti” znači tijelo koje provodi aktivnosti ocjenjivanja sukladnosti kao treća strana, uključujući testiranje, certifikaciju i inspekciju;

„prijavljeno tijelo” znači tijelo za ocjenjivanje sukladnosti prijavljeno u skladu s ovom Uredbom i drugim relevantnim zakonodavstvom Unije o usklađivanju;

„oznaka CE” znači oznaka kojom dobavljač označuje da je UI sustav sukladan sa zahtjevima utvrđenima u poglavlju III. odjeljku 2. i drugim primjenjivim zakonodavstvom Unije o usklađivanju kojim se propisuje stavljanje te oznake;

„tijelo za nadzor tržišta” znači nacionalno tijelo koje provodi aktivnosti i poduzima mjere na temelju Uredbe (EU) 2019/1020;
„dobavljač niže u lancu” znači dobavljač UI sustava, uključujući UI sustav opće namjene, u koji je integriran UI model, bez obzira na to dobavlja li UI model on sâm i je li vertikalno integriran ili ga dobavlja drugi subjekt na temelju ugovornih odnosa.

Na koje subjekte primjenjuje Uredba o umjetnoj inteligenciji? (članak 2. Uredbe o umjetnoj inteligenciji)

Uredba o umjetnoj inteligenciji primjenjuje se na javne i privatne subjekte unutar i izvan EU, sve dok je sustav umjetne inteligencije stavljen na tržište Unije ili njegova uporaba ima učinak na osobe koje se nalaze u EU-u.

Postoje određene iznimke od primjene Uredbe. Ova se Uredba se ne primjenjuje na UI sustave ili UI modele, uključujući njihove izlazne rezultate, koji su posebno razvijeni i stavljeni u upotrebu isključivo u svrhu znanstvenih istraživanja i razvoja. Također, ne primjenjuje se na aktivnosti istraživanja, testiranja ili razvoja UI sustava ili UI modela prije njihova stavljanja na tržište ili u upotrebu. Takve se aktivnosti provode u skladu s primjenjivim pravom Unije. Testiranje u stvarnim uvjetima nije obuhvaćeno tim isključenjem. Također, sustavi umjetne inteligencije koji su isključivo namijenjeni vojnim, obrambenim ili sigurnosnim svrhama države izuzeti su od primjene, neovisno o vrsti subjekta koji provodi te aktivnosti.

Primjenjuje li se Opća uredba o zaštiti podataka na obradu osobnih podataka prilikom obrade osobnih podataka u sustavima umjetne inteligencije?

Za svaku obradu osobnih podataka, pa tako i za obradu osobnih podataka u sustavima umjetne inteligencije, potrebno je prije svega odrediti odgovarajući pravni temelj i zakonitu svrhu.

Neovisno o tome kojeg je rizika sustav umjetne inteligencije, Opća uredba o zaštiti podataka se primjenjuje na aktivnosti obrade osobnih podataka u sustavima umjetne inteligencije tijekom cijelog životnog ciklusa sustava umjetne inteligencije; u svim fazama razvoja, testiranja i uporabe sustava.

Sukladno članku 47. Uredbe o umjetnoj inteligenciji, dobavljač sustava umjetne inteligencije (*AI provider*) je u obvezi sastaviti pisanu EU izjavu o sukladnosti iz članka 47. za visokorizični sustav. Ako UI sustav obuhvaća obradu osobnih podataka, u obvezi je dati izjavu da je taj UI sustav usklađen s Uredbama (EU) 2016/679 (Opća uredba o zaštiti podataka) i (EU) 2018/1725 te Direktivom (EU) 2016/680.

Životni ciklus sustava umjetne inteligencije, kako je definiran u standardima ISO/IEC 22989 i ISO/IEC 5338, pruža strukturirani okvir za razumijevanje toka podataka tijekom razvoja, implementacije i rada sustava umjetne inteligencije. Ovaj životni ciklus također je ključan za prepoznavanje i ublažavanje rizika za privatnost u svakoj fazi.

Izvor: *AI Privacy Risks & Mitigations: Large Language Models (LLMs)*

Jesu li dobavljači sustava umjetne inteligencije (*AI provideri*) ujedno i voditelji obrade u smislu Opće uredbe o zaštiti podataka?

Dobavljač UI sustava općenito će biti voditelj obrade za aktivnosti obrade koje je proveo radi razvoja UI sustava jer će dobavljač općenito definirati svrhu i sredstva takvih aktivnosti obrade. Na primjer, dobavljač koji pokreće razvoj UI sustava i odabire i stvara skup osobnih podataka za treniranje sustava umjetne inteligencije kojeg razvija smatrao bi se voditeljem obrade. To bi bio slučaj i kad bi dobavljač povjerio izradu takvog skupa podataka pružatelju usluga s pomoću dovoljno detaljnih dokumentiranih uputa (pri čemu potonji djeluje kao izvršitelj obrade).

Subjekt koji uvodi sustav umjetne inteligencije (*AI deployer*) općenito će biti voditelj obrade osobnih podataka u kontekstu uporabe tog sustava pod svojom nadležnošću jer će utvrditi zašto upotrebljava određeni UI sustav te koje podatke obrađuje. To će općenito biti slučaj kada subjekt koji uvodi UI sustav odluči upotrebljavati UI sustav u određenom kontekstu za određene ispitanike i određene kategorije osobnih podataka. Voditelj obrade je taj koji je u obvezi uskladiti sve aktivnosti obrade osobnih podataka u sustavima umjetne inteligencije s Općom uredbom o zaštiti podataka te mora biti u mogućnosti nadzornom tijelu (AZOP-u) dokazati usklađenost.

Primjer: Banka uvodi sustav umjetne inteligencije za automatsku procjenu kreditne sposobnosti klijenata. Ovaj sustav koristi osobne podatke kao što su prihodi, kreditna povijest, dob, status zaposlenja i ostali financijski pokazatelji kako bi predložio odluku o odobrenju ili odbijanju kredita. U ovom slučaju banka je subjekt koji uvodi sustav umjetne inteligencije (*deployer*) i istovremeno djeluje kao **voditelj obrade osobnih podataka** jer:

- određuje **svrhu** obrade (procjena kreditne sposobnosti klijenata),
- bira i određuje **sredstva obrade** (koji AI sustav koristiti, koji podaci su potrebni i kako će se koristiti),
- nadzire upotrebu sustava pod **svojom nadležnošću**.

Ti opći elementi i čimbenici ne bi trebali zamijeniti potrebu za analizom od slučaja do slučaja, posebno ako je za određenu aktivnost obrade uključeno više od jednog sudionika. Dobavljač sustava umjetne inteligencije mogao bi tijekom uvođenja UI sustava djelovati kao izvršitelj obrade podataka za određeni subjekt koji uvodi UI sustav ili kao voditelj obrade ako je ovlašten za obradu podataka o upotrebi za svoj račun (npr. za poboljšanje UI sustava ili modela koji pruža). Čak bi i subjekti za uvođenje mogli biti izvršitelji obrade ako obavljaju aktivnost obrade podataka u ime voditelja obrade, npr. ako svojim klijentima pružaju samo UI sustav bez ikakve kontrole nad time zašto se i kako upotrebljava.

Moguće je i da se dobavljač UI sustava i subjekt koji primjenjuje UI sustav smatraju zajedničkim voditeljima obrade, npr. ako subjekt koji primjenjuje UI sustav i dobavljač zajednički odrede koje će podatke o uporabi ponovno upotrijebiti za razvoj ili poboljšanje (novog) UI sustava.

Primjer: Softverska tvrtka razvila je sustav za pomoć pri dijagnostici, a bolnica je od softverske tvrtke kupila navedeno softversko rješenje. Sustav umjetne inteligencije obrađuje podatke pacijenata koje unosi bolnica, a softverska tvrtka nema pristup podacima osim u svrhu tehničkog održavanja sustava, sukladno ugovoru. U ovom slučaju softverska tvrtka djeluje kao izvršitelj obrade jer postupa isključivo prema uputama bolnice koja je voditelj obrade.

Primjer: *Startup* je izradio chatbot koji služi u obrazovne svrhe, a koristi ga škola. Tijekom uporabe chatbota, *startup* prikuplja i analizira podatke o interakcijama učenika sa sustavom kako bi poboljšao performanse modela. *Startup* određuje svrhu i način obrade tih podataka za vlastite potrebe i poboljšanje modela te je u ovom slučaju voditelj obrade.

Primjer: Softverska tvrtka razvila je alat umjetne inteligencije za automatsku analizu životopisa kandidata. Međutim ne koristi sustav za vlastite svrhe, već ga nudi kao uslugu velikim korporacijama. Softverska tvrtka ima nekoliko klijenata, velikih korporacija koje koriste softver i unose osobne podatke kandidata u sučelje alata te određuju samostalno sve parametre: svrhu obrade, kriterije i koje će kandidate pozvati na razgovor. Softverska tvrtka pruža samo tehničku podršku odnosno omogućuje da algoritam izvrši analizu na temelju parametara koje zadaje klijent, bez da sama odlučuje o tome kako i zašto se podaci obrađuju.

Primjer: Banka koristi sustav umjetne inteligencije za detekciju prijevara koji je izradio *AI startup*. Dogovoreno je da obje strane imaju pristup podacima o ponašanju korisnika, koje zatim zajednički koriste za treniranje nove verzije modela. Budući da zajednički određuju svrhu i sredstva obrade (npr. koje podatke koristiti i kako ih koristiti za daljnji razvoj), banka i *AI startup* su zajednički voditelji obrade.

U kontekstu razvoja sustava umjetne inteligencije, primjenjuje li se Opća uredba o zaštiti podataka na automatiziranu obradu bez sustava pohrane?

Opća uredba o zaštiti podataka se primjenjuje na obradu osobnih podataka koja se u cijelosti obavlja automatizirano te na neautomatiziranu obradu osobnih podataka koji čine dio sustava pohrane ili su namijenjeni biti dio sustava pohrane.

Primjer: Marketinška agencija je postavila interaktivne reklamne plakate na frekventnim gradskim lokacijama. Plakati sadrže ugrađene kamere koje koriste tehnologije umjetne inteligencije za analizu izraza lica prolaznika. Sustav kratkotrajno (npr. dvije sekunde) snima lice osobe radi procjene emocionalne reakcije na prikazani oglas (npr. sretan, zabrinut, nezainteresiran). Snimke se ne pohranjuju, ali se rezultat analize emocije bilježi i dalje koristi u svrhu evaluacije učinkovitosti kampanje. Snimke se ne pohranjuju, ali se rezultat analize emocije bilježi i dalje koristi u svrhu evaluacije učinkovitosti kampanje. Sukladno članku 2. stavku 1. Opće uredbe o zaštiti podataka, **Uredba primjenjuje na obradu osobnih podataka koja se u cijelosti obavlja automatizirano** te na neautomatiziranu obradu osobnih podataka koji čine dio sustava pohrane ili su namijenjeni biti dio sustava pohrane. U ovom slučaju radi se o automatiziranoj obradi osobnih podataka na koju se primjenjuje Opća uredba o zaštiti podataka iako se snimke pojedinaca ne pohranjuju.

Primjenjuje li se Opća uredba o zaštiti podataka na obradu pseudonimiziranih podataka u sustavima umjetne inteligencije?

„Pseudonimizacija" znači obrada osobnih podataka na način da se osobni podaci više ne mogu pripisati određenom ispitaniku bez uporabe dodatnih informacija, pod uvjetom da se takve dodatne informacije drže odvojeno te da podliježu tehničkim i organizacijskim mjerama kako bi se osiguralo da se osobni podaci ne mogu pripisati pojedincu čiji je identitet utvrđen ili se može utvrditi.

Pseudonimizirani podaci su i dalje osobni podaci te se Opća uredba o zaštiti podataka primjenjuje.

Primjer: Znanstveni institut razvija sustav umjetne inteligencije za analizu podataka o zdravlju korisnika radi otkrivanja rizika od razvoja određenih bolesti. U svrhu treniranja modela koriste se zdravstveni podaci pacijenata, pri čemu su imena, OIB-i i drugi izravni identifikatori uklonjeni, a svakom pacijentu dodijeljen je jedinstveni kod (npr. ID broj). Dodatna tablica s povezivanjem tih kodova s identitetima pacijenata čuva se odvojeno, uz stroge sigurnosne mjere pristupa. Osobni podaci pseudonimizirani, ali se identitet osoba može ponovno utvrditi uz pomoć dodatnih informacija, ti podaci se i dalje smatraju osobnim podacima prema Općoj uredbi o zaštiti podataka Pseudonimizacija je dodatna sigurnosna mjera koja smanjuje rizik, ali se Opća uredba o zaštiti podataka primjenjuje i na pseudonimizirane podatke.

Više o tehnikama pseudonimizacije možete saznati u Smjernicama o pseudonimizaciji Europskog odbora za zaštitu podataka

Kako su povezane Opća uredba o zaštiti podataka i Uredba o umjetnoj inteligenciji?

Uredbu o umjetnoj inteligenciji i Opću uredbu o zaštiti podataka treba smatrati kao komplementarne instrumente odnosno propise koji se međusobno nadopunjuju. Važno je naglasiti da se **pravo EU o zaštiti podataka u potpunosti primjenjuje na obradu osobnih podataka**

uključenih u životni ciklus sustava umjetne inteligencije, kako je izričito navedeno u članku 2. stavku 7. Uredbe o umjetnoj inteligenciji (vidjeti i uvodne izjave 9. i 10.).

Opća uredba o zaštiti podataka i Uredba o umjetnoj inteligenciji usko su povezane pravne norme, pri čemu je usklađenost s Općom uredbom o zaštiti podataka temeljni preduvjet za razvoj etičke, zakonite i antropocentrične umjetne inteligencije, koja je u skladu s temeljnim pravima, demokratskim načelima i vladavinom prava.

Ovi propisi ne predstavljaju prepreku inovacijama, već osiguravaju regulatorni okvir koji omogućuje razvoj umjetne inteligencije na način koji maksimizira društvenu korist, uz istodobno minimiziranje rizika.

Kako odrediti pravni temelj za obradu osobnih podataka u sustavima umjetne inteligencije?

Člankom 6. stavkom 1. Opće uredbe o zaštiti podataka propisano je da je obrada zakonita samo ako i u onoj mjeri u kojoj je ispunjeno najmanje jedno od sljedećega:

(a) ispitanik je dao privolu za obradu svojih osobnih podataka u jednu ili više posebnih svrha;

(b) obrada je nužna za izvršavanje ugovora u kojem je ispitanik stranka ili kako bi se poduzele radnje na zahtjev ispitanika prije sklapanja ugovora;

(c) obrada je nužna radi poštovanja pravnih obveza voditelja obrade;

(d) obrada je nužna kako bi se zaštitili životno važni interesi ispitanika ili druge fizičke osobe;

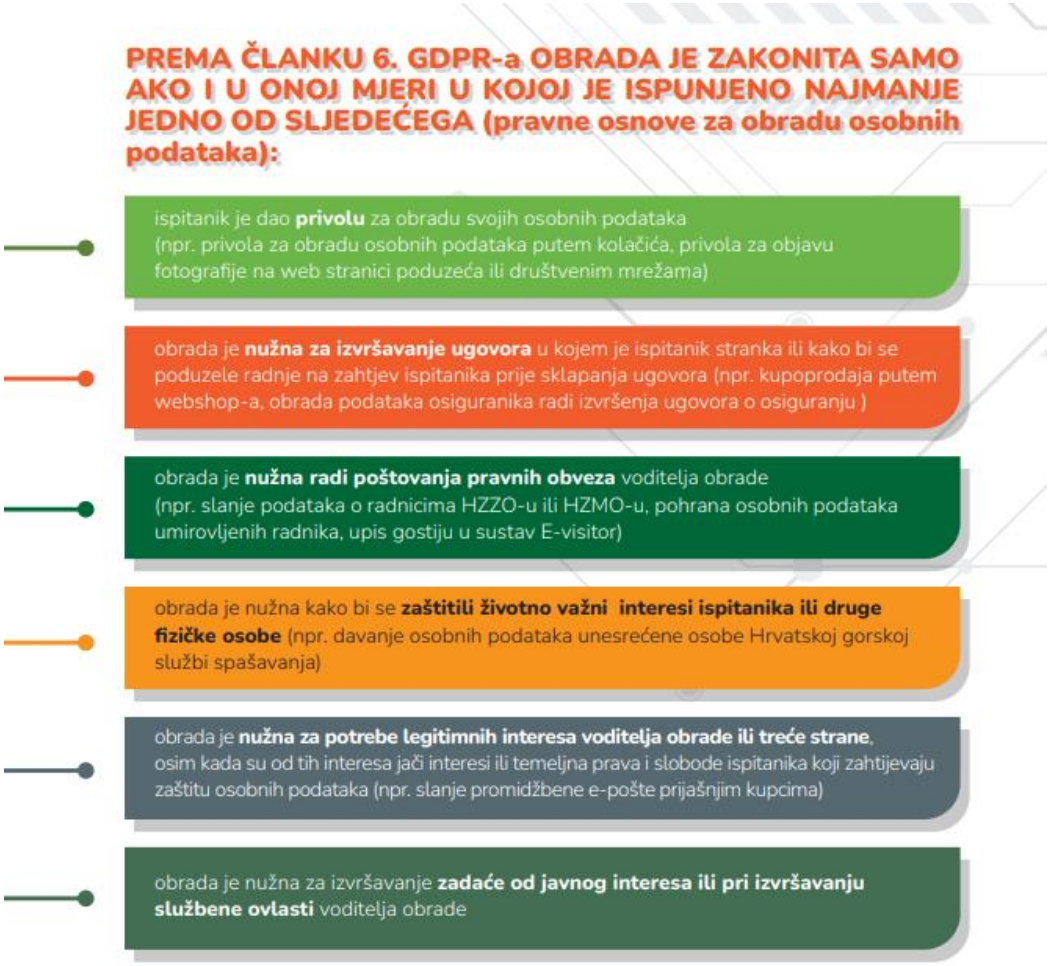
(e) obrada je nužna za izvršavanje zadaće od javnog interesa ili pri izvršavanju službene ovlasti voditelja obrade;

(f) obrada je nužna za potrebe legitimnih interesa voditelja obrade ili treće strane, osim kada su od tih interesa jači interesi ili temeljna prava i slobode ispitanika koji zahtijevaju zaštitu osobnih podataka, osobito ako je ispitanik dijete.

Točka (f) se ne odnosi na obradu koju provode tijela javne vlasti pri izvršavanju svojih zadaća.

Slijedom navedenog, da bi obrada osobnih podataka bila zakonita, za svaku obradu osobnih podataka učenika u sustavima umjetne inteligencije potrebno je prije svega identificirati odgovarajući pravni temelj.

U kontekstu razvoja i implementacije UI sustava najčešći pravni temelji bit će privola, ugovor i legitimni interes. Posebno je važno imati na umu da ispitanik mora biti u mogućnosti u svakom trenutku povući svoju privolu, pa samim time u mnogim situacijama privola neće biti odgovarajući pravni temelj za obradu osobnih podataka u sustavima umjetne inteligencije.



*Saznajte više o pravnim temeljima za obradu osobnih podataka na poveznici:

- Pravni temelji za obradu osobnih podataka

Može li legitimni interes biti pravni temelj za obradu osobnih podataka u svrhu treniranja modela umjetne inteligencije?

Kako bi se utvrdili može li se određena obrada osobnih podataka temeljiti na članku 6. stavku 1. točki (f) Opće uredbe, voditelji obrade moraju provesti trodijelni test (svrhe, nužnosti, ravnoteže) te isti dokumentirati kako bi nadzornom tijelu (AZOP-u) mogli dokazati da doista imaju legitimni interes za obradu osobnih podataka.

Pri tome je potrebno ispuniti tri kumulativna uvjeta: i. ostvarivanje legitimnog interesa od strane voditelja obrade ili treće strane; ii. obrada je nužna za ostvarivanje legitimnog interesa; i iii. legitimni interes nije podređen interesima ili temeljnim pravima i slobodama ispitanika.

Interes se može smatrati legitimnim ako su ispunjena sljedeća tri kumulativna kriterija:

a. interes je zakonit
b. interes je jasno i precizno artikuliran i
c. interes je stvaran i prisutan, a ne špekulativan.

Sljedeći primjeri mogu predstavljati legitiman interes u kontekstu UI modela:

i. razvoj usluge razgovornog agenta kako bi se pomoglo korisnicima;

ii. razvoj UI sustava za otkrivanje prijevarnog sadržaja ili ponašanja; i

iii. poboljšanje otkrivanja prijetnji u informacijskom sustavu.

Ovisno o slučaju, planiranu količinu osobnih podataka uključenih u UI model treba procijeniti imajući u vidu manje invazivne alternative koje bi razumno mogle biti dostupne kako bi se jednako učinkovito postigla svrha legitimnog interesa koji se nastoji ostvariti.

Ako je ostvarivanje te svrhe moguće i s pomoću UI modela koji ne podrazumijeva obradu osobnih podataka, smatrat će se da ta obrada osobnih podataka nije nužna. To je posebno važno za razvoj UI modela. Pri procjeni je li uvjet nužnosti ispunjen, AZOP će obratiti posebnu pozornost na količinu obrađenih osobnih podataka i je li ona proporcionalna za ostvarivanje predmetnog legitimnog interesa, također s obzirom na načelo smanjenja količine podataka.

Pri procjeni nužnosti trebalo bi uzeti u obzir i širi kontekst planirane obrade osobnih podataka. Postojanje sredstava koja su manje invanzivna za temeljna prava i slobode ispitanika može se razlikovati ovisno o tome je li voditelj obrade u izravnom odnosu s ispitanicima (podatci prve strane) ili ne (podatci trećih strana). Sud EU-a u predmetu C-621/22, Koninklijke Nederlandse Lawn Tennisbond (ECLI:EU:C:2024:857), točke 51.–53. naveo je neka razmatranja koja treba uzeti u obzir pri analizi nužnosti obrade podataka prve strane u svrhu legitimnih interesa koji se nastoje ostvariti (iako u kontekstu otkrivanja takvih podataka trećim stranama).

Provedba tehničkih zaštitnih mjera za zaštitu osobnih podataka također može doprinijeti ispunjavanju testa nužnosti. To bi moglo uključivati, na primjer, provedbene mjere na način da se ne postigne anonimizacija, ali kojima se i dalje smanjuje jednostavnost identifikacije ispitanika.

Treći korak procjene legitimnog interesa jest „testom ravnoteže“. Taj se korak sastoji od utvrđivanja i opisivanja različitih suprotstavljenih prava i interesa o kojima je riječ, tj. s jedne strane, interesa, temeljnih prava i sloboda ispitanika, a s druge strane interesa voditelja obrade ili treće strane. Potom bi trebalo razmotriti posebne okolnosti slučaja kako bi se dokazalo da je legitimni interes odgovarajuća pravna osnova za predmetne aktivnosti obrade.

Razumna očekivanja imaju ključnu ulogu u testu ravnoteže, među ostalim zbog složenosti tehnologije koja se upotrebljava u UI modelima i činjenice da bi ispitanicima moglo biti teško

razumjeti raznolikost mogućih uporaba UI modela i uključene obrade podataka. U tu se svrhu mogu razmotriti informacije pružene ispitanicima kako bi se procijenilo mogu li ispitanici razumno očekivati da će se njihovi osobni podatci obrađivati. Međutim, iako izostavljanje informacija može doprinijeti tome da ispitanici ne očekuju određenu obradu, samo ispunjenje zahtjeva u pogledu transparentnosti utvrđenih u Općoj uredbi nije dovoljno kako bi se smatralo da ispitanici mogu razumno očekivati određenu obradu . Nadalje, samo zato što su informacije koje se odnose na fazu razvoja UI modela uključene u politiku zaštite privatnosti voditelja obrade, to ne znači nužno da ispitanici mogu razumno očekivati da će se to dogoditi, nego bi to nadzorna tijela trebala analizirati s obzirom na posebne okolnosti slučaja i uzimajući u obzir sve relevantne čimbenike.

*Saznajte više na poveznicama:

- Mišljenje 28/2024 o određenim aspektima zaštite podataka povezanima s obradom osobnih podataka u kontekstu modela umjetne inteligencije Europskog odbora za zaštitu; Smjernice o legitimnom interesu; Test razmjernosti (legitimnog interesa)

Kako omogućiti UI-ju da "uči" i poboljša performanse, a istovremeno smanjiti količinu osobnih podataka koji se koriste?

Načela obrade osobnih podataka čine samu srž Opće uredbe o zaštiti podataka te prožimaju sve njezine odredbe. Kako bi obrada osobnih podataka bila zakonita, etička i poštena prema ispitanicima, voditelji obrade moraju se uvijek voditi načelima obrade osobnih podataka iz članka

5. Opće uredbe o zaštiti podataka. Navedeno je od iznimne važnosti u kontekstu uporabe osobnih podataka u sustavima umjetne inteligencije.

Osobni biti moraju biti zakonito, pošteno i transparentno obrađivani s obzirom na ispitanika (načelo zakonitosti, poštenosti i transparentnosti); prikupljeni u posebne, izričite i zakonite svrhe te se dalje ne smiju obrađivati na način koji nije u skladu s tim svrhama (načelo ograničavanja svrhe); primjereni, relevantni i ograničeni na ono što je nužno u odnosu na svrhe u koje se obrađuju (načelo smanjenja količine podataka); točni i prema potrebi ažurni (načelo točnosti); čuvani u obliku koji omogućuje identifikaciju ispitanika samo onoliko dugo koliko je potrebno u svrhe radi kojih se osobni podaci obrađuju (načelo ograničenja pohrane); i obrađivani na način kojim se osigurava odgovarajuća sigurnost osobnih podataka, uključujući zaštitu od neovlaštene ili nezakonite obrade te od slučajnog gubitka, uništenja ili oštećenja primjenom odgovarajućih tehničkih ili organizacijskih mjera (načelo cjelovitosti i povjerljivosti).

Voditelj obrade odgovara za usklađenost sa ovdje navedenim načelima Opće uredbe o zaštiti osobnih podataka te je mora biti u mogućnosti dokazati (načelo pouzdanosti).

Zakonitost, poštenost i transparentnost

U svakoj fazi razvoja, testiranja i uporabe umjetne inteligencije obrada osobnih podataka mora biti zakonita (potrebno je identificirati odgovarajući pravni temelj), te poštena i transparentna prema ispitanicima. To podrazumijeva da ispitanicima čiji se osobni podaci obrađuju takva obrada ne smije nanijeti nikakvu vrstu štete te da moraju biti obaviješteni o tome na koji način i u koje svrhe se obrađuju njihovi osobni podaci, koje su moguće posljedice i rizici, kako će isti biti ublaženi te koja su njihova prava.

Smanjenje količine podataka

Sustavi umjetne inteligencije, a osobito oni koji se temelje na metodama strojnog učenja, zahtijevaju velike količine podataka. Ti podaci ključni su ne samo za treniranje sustava, već i za njegovo testiranje, usporedbu i validaciju. Sastavljanje kvalitetnih skupova podataka podrazumijeva značajan trud, budući da podaci moraju biti pravilno označeni (anotirani), očišćeni, standardizirani i pripremljeni za uporabu.

Načelo smanjenja količine podataka propisuje da osobni podaci koji se prikupljaju i obrađuju moraju biti primjereni, relevantni i ograničeni na ono što je nužno u odnosu na jasno definiranu svrhu obrade. Posebna pažnja mora se posvetiti prirodi podataka, pri čemu ovo načelo zahtijeva još strožu primjenu kada je riječ o posebnim kategorijama osobnih podataka.

Iako uporaba velikih količina podataka može biti nužna za razvoj i funkcioniranje sustava umjetne inteligencije, primjena načela smanjenja količine podataka ne predstavlja prepreku takvoj obradi, već zahtijeva pažljivo planiranje i kontrolu.

U fazi učenja (treniranja algoritma) može biti prihvatljivo osigurati pristup većim količinama i raznolikim skupovima podataka, pod uvjetom da se primjenjuju odgovarajuće tehničke i organizacijske mjere koje su proporcionalne rizicima koje takva obrada nosi. Pri tome se osobito

mora voditi računa o prirodi podataka, njihovoj količini i svrsi razvoja sustava. Te mjere mogu uključivati, primjerice:

- ograničen pristup podacima na mali broj ovlaštenih osoba,
- obradu podataka u točno određenom vremenskom razdoblju,
- pseudonimizaciju podataka.

Nakon završetka faze učenja i prije prelaska u produkcijsku fazu (odnosno primjene sustava izvan kontroliranog, „laboratorijskog“ okruženja), potrebno je uvesti stroža pravila za nadzor obrade osobnih podataka. To uključuje ograničavanje obrade isključivo na one kategorije podataka koje su se tijekom faze razvoja pokazale nužnima za postizanje unaprijed određene svrhe, kao i definiranje dodatnih tehničkih i organizacijskih mjera zaštite. Važno je pritom uzeti u obzir da se zahtjevi za obradu podataka u produkcijskom okruženju razlikuju od onih u fazi dizajna i razvoja, osim u slučajevima kada i sama razvojna faza uključuje povišene rizike za prava i slobode ispitanika.

Ograničenje pohrane

Osobni podaci ne smiju se čuvati neograničeno. Opća uredba o zaštiti podataka zahtijeva da se unaprijed odredi razdoblje čuvanja podataka, nakon kojeg se podaci moraju izbrisati ili, u određenim slučajevima, arhivirati. Ovo razdoblje čuvanja mora utvrditi voditelj obrade, uzimajući u obzir svrhu zbog koje su osobni podaci prikupljeni.

U kontekstu implementacije sustava umjetne inteligencije, u pojedinim slučajevima može postojati potreba za duljim čuvanjem osobnih podataka u odnosu na druge oblike obrade. To je, primjerice, slučaj kod sastavljanja skupova podataka za treniranje modela i razvoj novih sustava, kao i radi ispunjavanja zahtjeva za sljedivost, evaluaciju učinkovitosti i mjerenje performansi sustava tijekom njegova korištenja u produkcijskom okruženju.

Međutim, potreba za definiranjem razdoblja čuvanja podataka ne predstavlja prepreku za obradu osobnih podataka u sustavima umjetne inteligencije, već osigurava da takva obrada bude zakonita i transparentna. Ovo razdoblje uvijek mora biti razmjerno svrsi obrade. Na primjer, ako je svrha duljeg čuvanja podataka praćenje performansi sustava, ta svrha mora biti jasno definirana i planirana, a podaci koji se čuvaju u tu svrhu moraju biti pažljivo odabrani.

Sama činjenica da se želi pratiti učinkovitost sustava tijekom vremena nije dovoljna da bi se opravdalo dugotrajno zadržavanje svih osobnih podataka. Potrebno je osigurati da se čuvaju isključivo oni podaci koji su za tu svrhu nužni.

Kako sustavi umjetne inteligencije mogu “učiti”, a da budu u skladu s Općom uredbom o zaštiti podataka (GDPR)?

Razvoj i treniranje sustava umjetne inteligencije često zahtijeva obradu velikih količina podataka, uključujući osobne podatke. Međutim, Opća uredba o zaštiti podataka propisuje jasna pravila kako

bi se osigurala zaštita prava pojedinaca tijekom takvih procesa. U nastavku su opisani neki od ključnih pristupa koji omogućuju razvoj UI sustava u skladu s načelima zaštite osobnih podataka.

Anonimizacija i pseudonimizacija podataka

Jedan od načina za zaštitu osobnih podataka i privatnosti tijekom treniranja UI sustava je korištenje anonimiziranih ili pseudonimiziranih podataka. Naime, GDPR se ne primjenjuje na anonimizirane podatke, što znači da se podaci koji su potpuno anonimizirani mogu koristiti bez ograničenja, jer nije moguće identificirati ispitanike.

S druge strane, pseudonimizacija podrazumijeva zamjenu osobnih identifikatora (primjerice imena) kodiranjem ili drugim metodama, čime se smanjuje rizik od identifikacije. Međutim, takvi podaci i dalje se smatraju osobnima, jer postoji mogućnost njihove ponovne identifikacije uz dodatne informacije.

Prikupljanje isključivo nužnih podataka

Prema načelu smanjenja količine podataka, osobni podaci smiju se prikupljati i obrađivati samo u mjeri u kojoj su nužni za ostvarenje jasno definirane svrhe. UI modeli se mogu dizajnirati i optimizirati tako da za svoj rad koriste minimalnu količinu podataka, čime se dodatno smanjuju rizici za privatnost. Primjerice, ako je svrha AI sustava analizirati koje proizvode korisnici najčešće kupuju, tada mu je dovoljan pristup informacijama o kategorijama proizvoda i vremenu kupnje. Nije potrebno obrađivati identifikacijske podatke poput IP adrese, imena ili datuma rođenja jer isti ne doprinose analizi navika, a povećavaju rizik za privatnost.

Brisanje podataka iz treniranih modela („Machine Unlearning“)

Pod određenim okolnostima, ispitanici imaju pravo tražiti brisanje svojih osobnih podataka. To uključuje i podatke koji su korišteni za treniranje UI modela.

Novi algoritmi omogućuju tzv. „machine unlearning“, odnosno selektivno uklanjanje podataka iz treniranih modela bez potrebe za potpunim ponovnim treniranjem cijelog modela. Time se omogućuje poštivanje prava ispitanika i nakon što su njihovi podaci već korišteni u procesu učenja.

*Saznajte više na poveznici:

- Smjernice vezano za ostvarivanje prava ispitanika u kontekstu umjetne inteligencije

Korištenje tehnologije za zaštitu privatnosti (Privacy Enhancing Technologies – PETs)

Tehnologije za zaštitu privatnosti (*eng. Privacy Enhancing Technologies* – PETs) ključni su alati za razvoj modela umjetne inteligencije koji omogućuju zaštitu privatnosti, prava intelektualnog vlasništva i osjetljivih informacija. Obuhvaćaju niz alata koji omogućuju sigurno prikupljanje, obradu i dijeljenje podataka, uz očuvanje privatnosti, povjerljivosti i sigurnosti informacija. Primjenjuju se u različitim kontekstima, uključujući razvoj i treniranje modela umjetne inteligencije, statističku analizu i suradnju između više strana bez razmjene sirovih podataka.

Primjeri takvih tehnologija:

Sigurna višestranačka računalna obrada (*eng. Secure Multiparty Computation* – SMPC) omogućuje više strana da zajednički obrađuju podatke bez međusobnog otkrivanja cjelovitih informacija, čime se osigurava zaštita podataka i minimizira rizik od povreda privatnosti.

Homomorfno šifriranje omogućuje izvođenje izračuna nad šifriranim podacima bez potrebe za njihovim dešifriranjem, što osigurava visoku razinu sigurnosti i povjerljivosti.

Sintetički podaci predstavljaju realistične, umjetno generirane skupove podataka koji se koriste kada pristup stvarnim podacima nije moguć zbog pravnih, etičkih ili tehničkih ograničenja.

Pouzdana okruženja za izvršavanje (*eng. Trusted Execution Environments* – TEE) omogućuju obradu podataka u izoliranom dijelu procesora, odvojenom od ostatka sustava, čime se smanjuje mogućnost neovlaštenog pristupa ili zlouporabe.

Dokazi bez otkrivanja (*eng. Zero-Knowledge Proofs)* omogućuju pojedincu da dokaže određenu informaciju (npr. da ispunjava neki uvjet) bez otkrivanja samog podatka, čime se omogućuje obrada podataka u skladu s načelom minimizacije.

Federativno učenje (*eng. Federated Learning)* je metoda koja omogućuje različitim subjektima da treniraju modele umjetne inteligencije na vlastitim, lokalnim podacima, bez potrebe za međusobnom razmjenom tih podataka. Svaki subjekt razvija svoj lokalni model, a zatim se dijelovi naučenih obrazaca (tzv. gradijenti) dijele i objedinjavanju u zajednički, globalni model koji je precizniji i učinkovitiji. Federativno učenje ima sličnosti s konceptom sigurne višestranačke računalne obrade, metode koja omogućuje više sudionika da zajednički obrade podatke i dobiju rezultat bez da otkrivaju vlastite podatke jedni drugima. Ipak, federativno učenje se ne smatra nužno vrstom te obrade.

Postoje dva osnovna pristupa federativnom učenju: centralizirani i decentralizirani.

Kod centraliziranog federativnog učenja, poslužitelj za koordinaciju razvija početni model koji se zatim šalje svim sudionicima (npr. bolnicama, tvrtkama, uređajima). Svaki od njih lokalno trenira model na vlastitim podacima te vraća rezultate treniranja (a ne same podatke). Koordinacijski poslužitelj potom prikuplja te rezultate, integrira ih i ažurira zajednički model. Taj se proces ponavlja kako bi se model postupno usavršavao. Ovakav pristup zahtijeva pouzdanog trećeg subjekta koji koordinira proces, što nije slučaj kod sigurne višestranačke računalne obrade.

U decentraliziranom pristupu ne postoji središnji poslužitelj. Umjesto toga, sudionici međusobno razmjenjuju rezultate treniranja i zajedno ažuriraju globalni model. Ovaj pristup ima prednost jer smanjuje sigurnosne rizike povezane s centraliziranim sustavima i uklanja mogućnost prekida rada zbog jednog neuspjelog sustava (tzv. jedinstvena točka kvara).

Primjer: više bolnica u različitim zemljama EU želi razviti UI model za rano prepoznavanje raka pluća na temelju rendgenskih snimki i medicinskih zapisa.

Federativno učenje omogućuje da:

1. Svaka bolnica zadržava podatke lokalno – rendgenske snimke i dijagnoze ostaju unutar njihovih IT sustava.
2. UI model se lokalno trenira u svakoj bolnici na temelju njihovih podataka.
3. Umjesto slanja slika i zapisa, bolnice šalju samo ažuriranja modela (npr. matematičke promjene u modelu).
4. Centralni poslužitelj kombinira ta ažuriranja u jedinstven, poboljšani model koji se zatim vraća svim bolnicama.

Diferencijalna privatnost (*eng. Differential Privacy)* predstavlja svojstvo skupa podataka ili baze podataka koje omogućuje formalno matematičko jamstvo da se podaci pojedinca ne mogu pouzdano izdvojiti ili prepoznati. Temelji se na principu dodavanja slučajnog šuma u podatke.

Središnji pojam diferencijalne privatnosti je **epsilon (ε)**, poznat i kao **parametar privatnosti**. On određuje količinu šuma koji se dodaje, odnosno razinu zaštite privatnosti.

Epsilon odražava maksimalnu količinu informacija koju treće strane mogu zaključiti o određenoj osobi, uključujući i to je li uopće sudjelovala u skupu podataka. Dodani šum omogućuje **vjerojatno poricanje**, tj. stvara uvjerljivu neizvjesnost o tome nalazi li se određena osoba u skupu podataka, čime se štiti njezina privatnost.

Primjena diferencijalne privatnosti može se provoditi na dva načina:

- **Interaktivni pristup (temeljen na upitima)** – šum se dodaje svakom odgovoru na upit, a mogućnost daljnjih upita prestaje kada se iskoristi cjelokupni budžet privatnosti (tj. kada postoji rizik od otkrivanja osobnih podataka).
- **Neinteraktivni pristup** – razina identifikabilnosti podataka određuje se unaprijed u skladu s postavljenim budžetom privatnosti. Ovaj pristup posebno je koristan za objavu agregiranih i anonimnih statističkih podataka.

Također, razlikujemo dvije vrste diferencijalne privatnosti:

- **Globalna diferencijalna privatnost** – šum se dodaje prilikom agregiranja podataka;
- **Lokalna diferencijalna privatnost** – šum se dodaje na razini pojedinačnih zapisa, prije nego što se podaci agregiraju, i to od strane samih korisnika.

Diferencijalna privatnost je metoda zaštite privatnosti koja se koristi kad organizacije žele objaviti statističke podatke, ali pritom ne žele otkriti podatke o pojedincima. Kako bi to postigli, u rezultate se namjerno unosi određena razina "buke", odnosno nasumičnih odstupanja. Ta odstupanja su dovoljno mala da ne naruše ukupnu korisnost podataka, ali su dovoljna da onemoguće identifikaciju konkretnih osoba.

*Saznajte više na poveznicama:

- Smjernice OECD-a o PETs-ovima; Rizici u velikim jezičnim modelima i mjere za njihovo ublažavanje; Federativno učenje

Imaju li pojedinci čiji osobni podaci služe za treniranje modela umjetne inteligencije prava iz Opće uredbe o zaštiti podataka?

Pojedinci imaju sljedeća prava u vezi sa svojim osobnim podacima:

- pravo na pristup (članak 15. GDPR-a)
- pravo na ispravak (članak 16. GDPR-a)
- pravo na brisanje (pravo na zaborav) (članak 17. GDPR-a)
- pravo na ograničenje obrade (članak 18. GDPR-a)
- pravo na prenosivost podataka kada je pravna osnova obrade privola ili ugovor (članak 20. GDPR-a)
- pravo na prigovor kada se obrada temelji na legitimnom interesu ili javnom interesu (članak 21. GDPR-a)
- pravo na povlačenje privole u bilo kojem trenutku kada se obrada temelji na privoli (članak 7. stavak 3. GDPR-a)

Ispitanici moraju moći ostvariti svoja prava kako na skupovima podataka za treniranje, tako i na samim modelima umjetne inteligencije, ako se ti modeli ne smatraju anonimnima.

Ostvarivanje prava treba biti moguće jednostavnim pisanim ili usmenim zahtjevom.

Voditelj obrade mora obavijestiti ispitanike o tome na koji način mogu ostvariti svoja prava i uspostaviti unutarnji postupak kojim se definiraju uvjeti za upravljanje ostvarivanjem prava i njihovo praćenje. U praksi je procedura odgovora na zahtjeve ispitanika razlikuje ovisno o tome odnose li se na podatke za treniranje ili na sam UI model.

Ako voditelj obrade više ne treba identificirati ispitanika i može dokazati da to nije moguće, ne mora pohranjivati ni prikupljati dodatne informacije samo radi omogućavanja ostvarivanja prava. U tom slučaju, prema članku 11. GDPR-a, mora o tome obavijestiti ispitanike, ako je moguće.

To je često slučaj kod stvaranja i korištenja skupova podataka za treniranje, bez obzira jesu li anotirani ili ne. Pružatelj AI sustava u pravilu ne mora identificirati osobe čiji se podaci nalaze u skupu za treniranje, niti mora zadržavati identifikatore u pseudonimiziranim skupovima samo radi naknadne identifikacije.

Ako voditelj obrade može identificirati osobe, dužan je odgovoriti na njihove zahtjeve za ostvarivanje prava, posebno u vezi sa pravom na pristup (članak 15. Opće uredbe o zaštiti podataka).

Ispitanici imaju pravo dobiti točne informacije o primateljima njihovih osobnih podataka i izvorima iz kojih je *dobavljač* prikupio osobne podatke. To im omogućuje ostvarivanje prava prema voditeljima obrade koji posjeduju njihove podatke, a što je posebno važno zbog složenog lanca aktera u razvoju AI sustava.

Informacija o primateljima ili kategorijama primatelja

Iako se informacije mogu ograničiti na kategorije primatelja, pravo na pristup omogućuje pojedincu da traži i konkretne identitete primatelja, kao što je potvrđeno u presudi Suda EU (C-154/21 od 12. siječnja 2023.).

Organizacija može pružiti samo kategorije primatelja ako nije moguće precizno ih identificirati ili ako je zahtjev očito neutemeljen ili pretjeran. Ako se skup podataka dijeli s velikim brojem trećih strana (npr. u istraživačke svrhe), preporučuje se uvođenje autentifikacijskog mehanizma ili API-ja za evidentiranje identiteta trećih strana i podataka kojima su pristupili.

Informacija o izvoru podataka

Ako podaci nisu izravno prikupljeni od ispitanika (npr. putem *data brokera*), ispitanik ima pravo znati njihov izvor.

Primjerice, kod ponovne upotrebe već dostupnih skupova podataka, mora se zabilježiti izvorni voditelj obrade i ispitaniku omogućiti kontakt s tim izvorom. Preporučuje se i opis sadržaja podataka i uvjeta njihova prikupljanja. Ako je skup podataka sastavljen iz više izvora, potrebno je osigurati sljedivost i pružiti informacije o svakom izvoru. Ako to nije moguće, pružaju se sve relevantne informacije o korištenim izvorima.

Primjer:
Tvrtka razvija AI model za analizu tržišnih trendova i koristi skup podataka kupljen od data brokera koji uključuje podatke o ponašanju potrošača. Podaci nisu prikupljeni izravno od ispitanika, već iz više online izvora i platformi.

Kako bi ispunila obveze iz članka 14. GDPR-a, tvrtka mora:

- **Zabilježiti tko je izvorni voditelj obrade**: npr. naziv i kontakt *data brokera*.
- **Omogućiti ispitaniku kontakt s tim izvorom**, kako bi mogao ostvariti svoja prava (npr. tražiti brisanje).
- **Opisati sadržaj skupa podataka** npr. da sadrži podatke o internetskim kupnjama, vremenu korištenja aplikacija i lokaciji uređaja.
- **Navesti uvjete pod kojima su podaci prikupljeni** npr. putem privole korisnika na partnerskim web-stranicama.

- **Osigurati sljedivost ako je korišteno više izvora**, primjerice označavanjem koje podatke je dostavio koji izvor.

Ako precizna identifikacija svih izvora nije moguća, ispitaniku se moraju pružiti **sve relevantne informacije o kategorijama izvora** npr. "internetske trgovine", "mobilne aplikacije", "društvene mreže", zajedno s općim podacima o tome kako su podaci sakupljeni.

Pravo na kopiju podataka korištenih za treniranje AI modela

Pravo na pristup uključuje i pravo na besplatnu kopiju svih obrađenih podataka. To uključuje i isječke iz skupova za treniranje, ako je to nužno za ostvarivanje drugih prava (presuda Suda EU, 4. svibnja 2023., C-487/21). Preporuka je dostaviti same podatke, ali i pripadajuće bilješke i metapodatke u razumljivom formatu. To ne smije ugroziti prava i slobode drugih, uključujući i prava drugih ispitanika, prava intelektualnog vlasništva ili poslovne tajne.

Pravo na ispravak, dopunu i brisanje

Ispitanici imaju pravo zatražiti ispravak netočnih ili nepotpunih podataka, uključujući netočne oznake (anotacije).

Pravo na brisanje uključuje situacije u kojima ispitanik obavijesti voditelja obrade da su korištene posebne kategorije osobnih podataka (u smislu članka 9. GDPR-a), a obrada nije bila opravdana iznimkom.

Obveza obavješćivanja o ispravku, brisanju i ograničenju obrade

Članak 19. GDPR-a propisuje da voditelj obrade mora obavijestiti sve primatelje kojima su podaci otkriveni o svakom ispravku, brisanju ili ograničenju, osim ako to nije moguće ili zahtijeva nerazmjeran napor.

To uključuje korištenje dostupnih tehnologija, a preporučuje se i upotreba API-ja ili barem sustava za bilježenje preuzimanja podataka.

Također, preporučuje se u ugovorima (npr. u licenci za ponovnu upotrebu skupa podataka) predvidjeti obvezu obavještavanja o ispravcima i brisanjima i za sve kasnije korisnike.

Generativni modeli i osobni podaci

Izlaz generativnog AI modela može predstavljati osobni podatak ako se odnosi na identificiranu ili osobu koju se može identificirati, neovisno o tome je li informacija točna. To je osobito relevantno u slučajevima tzv. regurgitacije sadržaja od strane velikih jezičnih modela (LLM), koji mogu proizvesti i stvarne i netočne informacije o pojedincima.

Međutim, ako se sadržaj generira isključivo na temelju statističke obrade prompta, a ne kao posljedica memoriranih podataka iz faze treniranja, tada odgovornost za obradu osobnih podataka snosi korisnik modela, a ne njegov dobavljač.

U iznimnim slučajevima model može proizvesti sintetske, izmišljene informacije koje slučajno odgovaraju stvarnim osobama. U takvim situacijama potrebno je provesti analizu u svakom konkretnom slučaju kako bi se utvrdilo primjenjuje li se Opća uredba o zaštiti podataka.

Identifikacija ispitanika unutar modela

Ako se na AI model primjenjuje Opća uredba o zaštiti podataka, voditelj obrade može, u skladu s člankom 11. GDPR-a, dokazati da ne može identificirati ispitanike. U tom slučaju nije dužan omogućiti ostvarivanje prava, osim ako ispitanik dostavi dodatne informacije koje omogućuju njegovu identifikaciju, a što bi voditelj obrade trebao omogućiti, ako je to izvedivo.

U većini slučajeva moguće je dokazati da trenutačne tehničke mogućnosti ne dopuštaju identifikaciju pojedinaca temeljem težina modela. Međutim, kod određenih vrsta modela poput onih koji eksplicitno pohranjuju podatke, primjerice support vector machines (SVM) ili algoritmi za grupiranje (klasteriranje) – tehnički je moguće povezati podatke s konkretnim osobama i time omogućiti ostvarivanje njihovih prava.

Primjer:
Dizajner velikog jezičnog modela (LLM), koji je treniran na podacima prikupljenima putem web scrapinga, može zatražiti od osobe da dostavi URL konkretne stranice i dio sadržaja koji želi provjeriti. Ako stranica više nije dostupna, može zatražiti izravan tekst za provjeru. Na temelju tog teksta moguće je provesti test regurgitacije kako bi se utvrdilo je li sadržaj bio dio skupa za treniranje.

Saznajte više na poveznici:

- Smjernice vezano za ostvarivanje prava ispitanika u kontekstu umjetne inteligencije

Kad se model umjetne inteligencije smatra anonimnim?

Kako bi se model mogao smatrati anonimnim, potrebno je ispuniti sljedeće uvjete:

1. Vjerojatnost izravnog izdvajanja osobnih podataka osoba čiji su podaci korišteni za treniranje modela mora biti beznačajna.
2. Vjerojatnost pribavljanja osobnih podataka putem modela bilo namjerno ili nenamjerno, putem upita također mora biti beznačajna.

Ove vjerojatnosti treba procjenjivati uzimajući u obzir sva sredstva za koja je razumno očekivati da će ih voditelj obrade ili bilo koja druga osoba koristiti, uključujući mogućnost nenamjerne (ponovne) upotrebe ili otkrivanja modela.

Na procjenu utječe i kontekst korištenja modela, primjerice, radi li se o internom modelu koji koriste samo zaposlenici, ili o modelu dostupnom široj javnosti.

U cilju smanjenja rizika, preporučuju se sljedeće mjere:

- Pseudonimizacija i filtriranje podataka tijekom pripreme skupa za treniranje
- Korištenje *Privacy Enhancement* Tehnologija (PETs)
- Smanjenje količine osobnih podataka koji se koriste u treniranju
- Testiranje modela u svrhu otkrivanja ranjivosti i provjere otpornosti na napade
- Tehničke mjere za sprječavanje napada na otkrivanje korištenih podataka za treniranje, poput napada na članstvo (eng. *membership inference attack*) i eksfiltraciju podataka

Saznajte više na poveznicama: Mišljenje 28/2024 o određenim aspektima zaštite podataka povezanima s obradom osobnih podataka u kontekstu modela umjetne inteligencije Europskog odbora za zaštitu; Usklađenost, zaštita osobnih podataka i sigurnost u UI sustavima; Osnove sigurnosti osobnih podataka u UI sustavima

Ako su osobni podaci javno dostupni, primjenjuje li se i na njih Opća uredba o zaštiti podataka?

Opća uredba o zaštiti podataka primjenjuje se i na osobne podatke koji su javno dostupni, primjerice objavljeni u javnim registrima, društvenim mrežama, internetskim stranicama javnih tijela itd. Ukoliko voditelj obrade (npr. *dobavljač*) želi koristiti takve osobne podatke za treniranje AI modela ili u neku drugu svrhu potrebno je odrediti odgovarajući pravni temelj, poštovati načela iz članka 5. Opće uredbe o zaštiti podataka i udovoljiti svim ostalim zahtjevima iz Opće uredbe o zaštiti podataka. Ukoliko voditelj obrade želi obradu osobnih podataka temeljiti na legitimnom interesu, takav legitiman interes je potrebno dokazati provedbom trodijelnog testa

Saznajte više na poveznicama: Rješenje AZOP-a o upravnoj novčanoj kazni zbog obrade osobnih podataka bez dokazanog pravno temelja; Novčana kazna izrečena *Clearviewu*

Smiju li se sustavi umjetne inteligencije koristiti za donošenje automatiziranih odluka koje utječu na pojedince?

Organizacije moraju osigurati usklađenost s člankom 22. Opće uredbe o zaštiti podataka (GDPR) kada se osobni podaci obrađuju u kontekstu korištenja UI sustava za donošenje automatiziranih odluka. Članak 22. štiti pojedince od odluka koje se temelje isključivo na automatiziranoj obradi, uključujući profiliranje, a koje proizvode pravne učinke za ispitanika ili na sličan način značajno utječu na njega.

Ispitanik ima pravo da se na njega ne odnosi odluka koja se temelji isključivo na automatiziranoj obradi, uključujući izradu profila, koja proizvodi pravne učinke koji se na njega odnose ili na sličan način značajno na njega utječu.

U uvodnoj izjavi 71. Opće uredbe o zaštiti podataka kao primjeri takvih odluka navode se, primjerice, automatsko odbijanje zahtjeva za kredit podnesenog putem interneta ili automatizirani postupci zapošljavanja bez ikakvog ljudskog uključivanja.

Automatizirano donošenje odluka je dozvoljeno, ako je odluka:

(a) potrebna za sklapanje ili izvršenje ugovora između ispitanika i voditelja obrade podataka;

(b) dopuštena pravom Unije ili pravom države članice kojem podliježe voditelj obrade te koje također propisuje odgovarajuće mjere zaštite prava i sloboda te legitimnih interesa ispitanika; ili

(c) temeljena na izričitoj privoli ispitanika.

Čak i u tim slučajevima, ispitanik mora biti obaviješten o tome da se odluka donosi automatizirano, uz obrazloženje logike koja stoji iza takve odluke te informaciju o njezinim posljedicama za ispitanika. Osim toga, osim ako za odluku ne postoji odgovarajuća pravna osnova, ispitanik ima pravo osporiti odluku, izraziti svoje stajalište i zatražiti uključivanje ljudskog faktora u njezino preispitivanje.

Jesu li organizacije ili pojedinci obvezni poštovati GDPR i ako nisu sami razvili UI sustav?

Svaka fizička ili pravna osoba koja određuje svrhu i sredstva obrade osobnih podataka smatra se voditeljem obrade te je obvezna postupati u skladu s GDPR-om, bez obzira na to je li sama razvila UI sustav ili koristi sustav treće strane. Čak i ako dobavljač ili subjekti koji uvode sustav određuje tehničke specifikacije sustava, to ne mijenja činjenicu da je upravo korisnik UI sustava najčešće voditelj obrade s odgovornostima prema GDPR-u.

Na što organizacije trebaju obratiti pozornost pri korištenju UI sustava trećih strana?

Organizacije moraju procijeniti postoji li prilikom korištenja „stranih" sustava rizik da će osobni podaci biti preneseni proizvođaču sustava ili drugim trećim stranama, što bi moglo dovesti do neovlaštenog otkrivanja poslovnih tajni ili povjerljivih podataka. Odnosno, moraju dobro razumjeti arhitekturu samog UI sustava i na koji način se odvija tok podataka unutar istoga i kojim podacima ima pristup. Uz to svakako je potrebno adresirati pitanje sigurnosti same obrade s obzirom na činjenicu da su UI sustavi nerijetko ranjivi na različite metode tradicionalnih kibernetičkih napada, ali i na nove.

Kako bi se rizici ublažili, potrebno je provesti procjenu konkretne situacije i uspostaviti interne smjernice koje će jasno definirati koje vrste podataka se smiju obrađivati korištenjem takvih sustava. U slučaju nejasnoća, preporučuje se konzultirati dobavljača sustava prije početka korištenja.

Mnogi pružatelji usluga nude i tzv. *on-premise* rješenja, koja omogućuju da se svi podaci obrađuju i pohranjuju isključivo na serverima same organizacije, čime se dodatno smanjuju rizici povezani s prijenosom podataka trećim stranama.

Je li moguće obrađivati posebne kategorije osobnih podataka kako bi se spriječila diskriminacija u UI sustavima?

U skladu s **člankom 10. stavkom 5.** Akta o umjetnoj inteligenciji, u mjeri u **kojoj je to nužno kako bi se osiguralo otkrivanje i ispravljanje pristranosti visokorizičnih UI sustava,** dobavljači tih visokorizičnih sustava iznimno mogu **obrađivati** posebne kategorije osobnih podataka, **podložno odgovarajućim zaštitnim mjerama u pogledu** temeljnih prava i sloboda pojedinaca. U istom se članku ta iznimna mogućnost podvrgava nizu kumulativnih uvjeta.

Člankom 10. stavkom 5. Akta o umjetnoj inteligenciji doista su predviđeni posebni zahtjevi u mjeri u kojoj je obrada dopuštena samo ako:

(a) cilj otkrivanja i ispravljanja pristranosti ne može se učinkovito ostvariti bez upotrebe posebnih kategorija osobnih podataka,
(b) obrada podliježe tehničkim ograničenjima za ponovnu uporabu osobnih podataka i najsuvremenijim mjerama sigurnosti i zaštite privatnosti, na primjer pseudonimizaciji,
(c) obrada podliježe odgovarajućim tehničkim i organizacijskim mjerama, na primjer odgovarajućoj kontroli pristupa,
(d) druge strane ne prenose i ne prenose osobne podatke niti im na drugi način pristupaju,
(e) posebne kategorije osobnih podataka brišu se nakon što se pristranost ispravi ili osobni podaci dosegnu kraj razdoblja njihova zadržavanja i
(f) evidencija aktivnosti obrade sadržava razloge zbog kojih je obrada posebnih kategorija osobnih podataka bila strogo nužna za otkrivanje i ispravljanje pristranosti te zašto se taj cilj ne bi mogao postići obradom drugih podataka.

Međutim, pružatelj usluga mora procijeniti jesu li ti zahtjevi ispunjeni, uz sve druge zahtjeve iz Opće uredbe o zaštiti podataka, te je odgovoran za procjenu (usp. članak 5. stavak 2. Opće uredbe o zaštiti podataka).

Za obradu posebnih kategorija osobnih podataka potrebno je odrediti pravni temelj iz članka 6. GDPR-a i iznimku iz članka 9. Opće uredbe o zaštiti podataka.

Dakle, kako bi se UI modeli mogli trenirati na posebnim kategorijama osobnih podataka u svrhu sprječavanja diskriminacije, potrebno je odrediti pravni temelj iz članka 6. stavka 1. Opće uredbe o zaštiti podataka i iznimku iz članka 9. Opće uredbe o zaštiti podataka, uz napomenu da **članak 10. stavak 5. UI Akta dozvoljava obradu posebnih kategorija osobnih podataka pod točno određenim uvjetima u svrhu sprječavanja diskriminacije te uz poduzimanje zaštitnih mjera, a primjenjuje se isključivo na visokorizične sustave**.

Kako se manifestira "*pristup temeljen na riziku*" na kojem se temelji Uredba o umjetnoj inteligenciji? Što to konkretno znači, kako se kategoriziraju različite razine rizika?

Pristup temeljen na riziku znači da se regulacija umjetne inteligencije prilagođava razini rizika koji određeni sustavi umjetne inteligencije predstavlja za društvo, temeljna prava i sigurnost. UI sustavi se kategoriziraju u četiri razine rizika: **neprihvatljiv rizik, visok rizik, ograničeni rizik i minimalni (niski) rizik**. Ovisno o kategoriji, primjenjuju se različite regulatorne obveze i ograničenja.

Uredba o umjetnoj inteligenciji temelji se na pristupu utemeljenom na riziku, a što znači da će opseg obveza za dobavljače *(provider*) i subjekte koji uvode sustave umjetne inteligencije (*deployer*) ovisiti o razini rizika koju konkretni sustav predstavlja za prava i sigurnost pojedinaca.

Neprihvatljiv rizik (zabranjeni sustavi)

Svi UI sustavi koji se smatraju jasnom prijetnjom za sigurnost, egzistenciju i prava ljudi zabranjeni su. Uredba o umjetnoj inteligenciji zabranjuje osam praksi, i to:

- štetnu manipulaciju i obmanu temeljenu na UI-u
- štetno iskorištavanje ranjivosti pomoću UI-a
- društveno ocjenjivanje (*eng. social scoring*)
- procjenu ili predviđanje individualnog kaznenog ponašanja
- neselektivno prikupljanje podataka s interneta ili CCTV snimki za stvaranje ili proširivanje baza podataka za prepoznavanje lica
- prepoznavanje emocija na radnom mjestu i u obrazovnim ustanovama
- biometrijsku kategorizaciju radi zaključivanja o zaštićenim osobinama (npr. rasi, spolu, političkom uvjerenju)
- **biometrijsku identifikaciju u stvarnom vremenu** na daljinu za potrebe provedbe zakona na javno dostupnim mjestima

*Saznajte više na poveznici: Smjernice Europske komisije o zabranjenim praksama u području umjetne inteligencije

Visok rizik (strogo regulirani sustavi)

Upotrebe UI-a koje mogu predstavljati ozbiljne rizike za zdravlje, sigurnost ili temeljna prava klasificiraju se kao visokorizične. Među njih spadaju:

- UI sigurnosne komponente u kritičnoj infrastrukturi (npr. promet), čiji kvar može ugroziti život i zdravlje građana;
- UI rješenja u obrazovanju, koja mogu utjecati na pristup obrazovanju i profesionalni put osobe (npr. ocjenjivanje ispita);
- UI sigurnosne komponente proizvoda (npr. UI primjene u robotski potpomognutim operacijama);
- UI alati za zapošljavanje, upravljanje radnicima i pristup samozapošljavanju (npr. AI alat za sortiranje životopisa);
- UI sustavi koji odlučuju o pristupu osnovnim privatnim i javnim uslugama (npr. kreditno bodovanje koje može utjecati na dobivanje zajma);
- UI sustavi za biometrijsku identifikaciju na daljinu, prepoznavanje emocija i biometrijsku kategorizaciju (npr. retroaktivno prepoznavanje počinitelja krađe);
- UI primjene u provedbi zakona koje mogu zadirati u temeljna prava (npr. procjena vjerodostojnosti dokaza);
- UI sustavi u upravljanju migracijama, azilom i nadzorom granica (npr. automatizirana provjera zahtjeva za vizu);
- UI rješenja u pravosuđu i demokratskim procesima (npr. UI alati za pripremu sudskih odluka).

UI sustavi visokog rizika podliježu strogim obvezama prije nego što se mogu staviti na tržište, uključujući:

- odgovarajuće sustave procjene i ublažavanja rizika;
- kvalitetne skupove podataka kako bi se smanjio rizik od diskriminacije;
- bilježenje aktivnosti radi osiguravanja sljedivosti rezultata;
- detaljnu dokumentaciju koja omogućuje nadležnim tijelima procjenu usklađenosti;
- jasne i dostatne informacije za korisnika UI sustava;
- odgovarajuće mjere ljudskog nadzora;
- visoku razinu robusnosti, kibernetičke sigurnosti i točnosti.

Ograničeni rizik (zahtjevi transparentnosti i tehnička dokumentacija)

U ovu kategoriju spadaju sustavi umjetne inteligencije (UI) koji mogu utjecati na ponašanje korisnika ili oblikovati njihove odluke, ali ne predstavljaju visok rizik za njihova temeljna prava ili sigurnost. Glavni regulatorni zahtjevi uključuju transparentnost i tehničku dokumentaciju.

To znači da korisnici moraju biti jasno informirani da komuniciraju sa sustavom umjetne inteligencije ili da konzumiraju sadržaj generiran putem iste.

Primjeri:

Chatbotovi: korisnici moraju znati da ne komuniciraju s čovjekom, već s UI sustavom.

***Deepfake* sadržaji**: slike, videozapisi i drugi generirani sadržaji moraju biti jasno označeni kao umjetno stvoreni.

UI u sustavima za preporuke: npr. personalizirani oglasi koji koriste UI analizu; korisnici moraju biti obaviješteni da je sadržaj rezultat UI obrade.

Minimalan ili nikakav rizik

Uredba o umjetnoj inteligenciji ne propisuje posebna pravila za UI sustave koji se smatraju niskorizičnima ili bez rizika. Velika većina UI sustava koji se trenutno koriste u EU spada u ovu kategoriju. To uključuje, primjerice, UI u videoigrama ili filtre za neželjenu poštu (spam).

Većina sustava umjetne inteligencije može se razvijati i koristiti u skladu s postojećim zakonodavstvom, bez dodatnih pravnih obveza. Dobavljači takvih sustava mogu se, na dobrovoljnoj osnovi, odlučiti primjenjivati zahtjeve za pouzdanu umjetnu inteligenciju i pridržavati se dobrovoljnih kodeksa ponašanja.

Osim toga, Uredba o umjetnoj inteligenciji uzima u obzir i sistemske rizike koji mogu proizaći iz modela umjetne inteligencije opće namjene, uključujući velike generativne modele umjetne inteligencije. Ti se modeli mogu koristiti za različite zadatke i sve češće predstavljaju temelj za mnoge sustave umjetne inteligencije u EU-u. Neki od tih modela mogli bi predstavljati sistemski rizik ako su iznimno sposobni ili široko primjenjivani.

Na primjer, snažni modeli mogli bi uzrokovati ozbiljne nesreće ili biti zloupotrijebljeni za provođenje opsežnih kibernetičkih napada. Također, veliki broj osoba mogao bi biti pogođen ako model širi štetne pristranosti kroz brojne različite primjene.

*Rizik povezan s transparentnošću

Ovaj rizik odnosi se na potrebu za transparentnošću u vezi s uporabom UI-a. Uredba o umjetnoj inteligenciji uvodi posebne obveze informiranja kako bi se očuvalo povjerenje ljudi. Na primjer,

pri korištenju UI sustava poput chatbotova, ljudi moraju biti informirani da komuniciraju sa strojem kako bi mogli donijeti informiranu odluku.

Pojedinci moraju biti u informirani i moraju biti u mogućnosti jasno prepoznati da je određeni sadržaj generirala umjetna inteligencija. Osim toga, određeni UI-generirani sadržaji moraju biti jasno i vidljivo označeni – osobito *deepfake sadržaji* i tekstovi koji imaju svrhu informiranja javnosti o temama od javnog interesa.

Primjenjuje li se Uredba o umjetnoj inteligenciji na generativnu umjetnu inteligenciju?

UI modeli opće namjene igraju značajnu ulogu u poticanju inovacija i širenju uporabe UI-a unutar Europske unije, jer se mogu koristiti za različite zadatke i integrirati u širok raspon UI sustava.

Zbog toga dobavljači takvih modela imaju određene obveze prema Uredbi o umjetnoj inteligenciji.

Veliki generativni UI modeli tipičan su primjer UI modela opće namjene s obzirom na to da omogućuju fleksibilnu proizvodnju sadržaja, primjerice u obliku teksta, audiozapisa, slika ili videozapisa, koji mogu lako obavljati širok raspon različitih zadaća.

Te obveze uključuju davanje informacija dobavljačima UI sustava koji namjeravaju integrirati model u svoje sustave te uspostavu politike usklađene s europskim zakonodavstvom o autorskom pravu. Osim toga, dobavljači najnaprednijih ili najutjecajnijih općih UI modela, odnosno onih koji predstavljaju sustavne rizike, podliježu dodatnim obvezama procjene i ublažavanja tih sustavnih rizika. Sustavni rizici uključuju rizike za temeljna prava i sigurnost te rizike povezane s gubitkom kontrole nad modelom.

*Saznajte više na poveznicama: Smjernice Europske komisije o obvezama dobavljača UI modela opće namjene; Kodeks dobre prakse za UI modele opće namjene; Predložak za pružatelje općih modela umjetne inteligencije za sažetak podataka korištenih za treniranje modela

Opći modeli umjetne inteligencije: UI model smatra se općim modelom ako je treniran koristeći količinu računalnih resursa (tzv. „compute") veću od 10^23 operacija nad brojevima s pomičnim

zarezom u sekundi te ako može generirati jezik (u obliku teksta ili zvuka), tekst u sliku ili tekst u video.

Pružatelji općih UI modela: Smjernice objašnjavaju pojmove „dobavljač" i „stavljanje na tržište" te pojašnjavaju kada se subjekt koji modificira opći UI model smatra dobavljačem.

Izuzeća od određenih obveza: Smjernice pojašnjavaju pod kojim uvjetima dobavljači općih UI modela koji su objavljeni pod besplatnom i otvorenom licencom, a ispunjavaju određene uvjete transparentnosti, mogu biti izuzeti od pojedinih obveza prema Uredbi o umjetnoj inteligenciji.

Provedba obveza: Smjernice pojašnjavaju implikacije za dobavljače općih UI modela koji se odluče pridržavati Kodeksa dobre prakse za opće UI modele te iznose očekivanja Komisije u pogledu usklađenosti od 2. kolovoza 2025.

Obveze iz Uredbe o umjetnoj inteligenciji za dobavljače općih UI modela počinju se primjenjivati **2. kolovoza 2025.** Od tog datuma nadalje, dobavljači koji stavljaju opće UI modele na tržište moraju poštivati svoje obveze prema Uredbi. Dobavljači općih UI modela koji će se klasificirati kao modeli s sustavnim rizikom moraju bez odgode obavijestiti Ured za umjetnu inteligenciju. Tijekom prve godine primjene obveza, Ured će blisko surađivati s dobavljačima, osobito s onima koji se pridržavaju Kodeksa dobre prakse, kako bi im pomogao u usklađivanju s pravilima. **Od 2. kolovoza 2026. Komisija počinje s primjenom svojih ovlasti za nadzor i provedbu.**

Dobavljači općih UI modela koji su već dostupni na tržištu prije 2. kolovoza 2025. moraju ispuniti relevantne obveze iz Uredbe o umjetnoj inteligenciji do 2. kolovoza 2027.

Koje su obveze za dobavljače sustava umjetne inteligencije (*eng. provider*)?

Prije stavljanja visokorizičnog UI sustava na tržište EU-a ili njegove uporabe, **dobavljači moraju provesti ocjenu sukladnosti (conformity assessment).** Time dokazuju da njihov sustav udovoljava obveznim zahtjevima za pouzdanu umjetnu inteligenciju, poput:

- kvalitete podataka,
- dokumentacije i sljedivosti,
- transparentnosti,
- ljudskog nadzora,
- točnosti,
- kibernetičke sigurnosti
- i robusnosti.

Ocjena sukladnosti mora se ponoviti ako dođe do značajnih promjena sustava ili njegove svrhe.

UI sustavi koji služe kao sigurnosne komponente proizvoda obuhvaćenih sektorskim zakonodavstvom EU-a uvijek će se smatrati visokorizičnima ako podliježu ocjeni sukladnosti treće

strane prema tom zakonodavstvu. Osim toga, svi biometrijski sustavi, bez obzira na njihovu primjenu, moraju proći ocjenu sukladnosti treće strane.

Pružatelji visokorizičnih UI sustava moraju također uspostaviti sustave upravljanja kvalitetom i rizicima kako bi osigurali usklađenost s novim zahtjevima i smanjili rizike za korisnike i osobe na koje sustav utječe, čak i nakon što je proizvod stavljen na tržište.

Visokorizični UI sustavi koje koriste javna tijela ili subjekti koji djeluju u njihovo ime moraju biti registrirani u javnu EU bazu podataka, osim ako se koriste za provedbu zakona i upravljanje migracijama; tada će se registrirati u ne-javni dio baze kojem mogu pristupiti samo nadležna nadzorna tijela.Kako bi se osigurala usklađenost tijekom cijelog životnog ciklusa UI sustava, nadzorna tijela će provoditi redovite revizije, nadgledati situaciju nakon puštanja na tržište i omogućiti pružateljima da dobrovoljno prijave ozbiljne incidente ili kršenja temeljnih prava ako ih uoče. U iznimnim slučajevima, nadležna tijela mogu odobriti izuzeća za pojedine visokorizične UI sustave da se stave na tržište.

U slučaju kršenja propisa, ova pravila omogućuju nacionalnim tijelima pristup informacijama potrebnima za istragu o tome je li uporaba UI sustava bila u skladu sa zakonom.

Procjena sukladnosti prije stavljanja na tržište (članci 43.- 47. Uredbe o umjetnoj inteligenciji).

Prije nego što visokorizični UI sustav bude stavljen na tržište ili u uporabu, dobavljač sustava umjetne mora provesti postupak ocjene sukladnosti kako bi dokazao da je sustav siguran i u skladu s odredbama Uredbe o umjetnoj inteligenciji.

Tehnička dokumentacija i registracija u EU bazi (članak 9. Uredbe o umjetnoj inteligenciji)

Dobavljači sustava visokorizične umjetne inteligencije su u obvezi:

- uspostaviti i održavati sustav upravljanja rizikom

- podaci i upravljanje podacima: skupovi podataka za treniranje, validaciju i testiranje moraju biti relevantni, dovoljno reprezentativni te, u najvećoj mogućoj mjeri, bez pogrešaka i potpuni s obzirom na namjenu. Moraju imati odgovarajuća statistička obilježja, među ostalim u odnosu na osobe ili skupine osoba u pogledu kojih se visokorizični UI sustav namjerava koristiti, ako je primjenjivo. Te se karakteristike skupova podataka mogu postići na razini pojedinačnih skupova podataka ili na razini kombinacije skupova podataka (članak 10. Uredbe o umjetnoj inteligenciji)

- voditi i redovito ažurirati tehničku dokumentaciju koja dokazuje sukladnost sustava s regulatornim zahtjevima (članak 11. Uredbe o umjetnoj inteligenciji)

- tehnički omogućiti automatsko evidentiranje događaja („evidencija”) tijekom cijelog životnog ciklusa sustava (članak 12. Uredbe o umjetnoj inteligenciji)

- uspostaviti sustav upravljanja kvalitetom (članak 17. Uredbe o umjetnoj inteligenciji)

- registrirati visokorizični UI sustav u posebnu EU bazu prije njegova stavljanja na tržište (članak 49. Uredbe o umjetnoj inteligenciji).

Transparentnost i objašnjivost (članak 13., članak 16. stavak h) Uredbe o umjetnoj inteligenciji)

Visokorizični UI sustavi moraju biti osmišljeni tako da:

-omogućuju razumljivo objašnjenje funkcionalnosti i odluka sustava krajnjim korisnicima i nadzornim tijelima, posebno u područjima poput zapošljavanja, obrazovanja, kreditne sposobnosti i pravosuđa, gdje je objašnjivost presudna za zaštitu prava ispitanika;

- korisnicima i nadzornim tijelima moraju biti pružene razumljive informacije.

Ljudski nadzor (članak 14. Uredbe o umjetnoj inteligenciji)

U radu visokorizičnih sustava mora biti osiguran učinkovit ljudski nadzor kako bi se spriječile štetne posljedice:

- ključne odluke ne smije donositi UI bez mogućnosti ljudske intervencije, osobito u područjima kao što su medicinska dijagnostika, zapošljavanje ili socijalna skrb.

Kibernetička sigurnost i otpornost na napade (članak 15. Uredbe o umjetnoj inteligenciji)

Visokorizični sustavi moraju:

- biti otporni na zlouporabu, manipulacije i kibernetičke napade,

-uključivati mehanizme za otkrivanje anomalija i osiguranje integriteta podataka

- visokorizični UI sustavi moraju biti otporni što je više moguće na greške, kvarove ili nedosljednosti koji se mogu dogoditi unutar sustava ili okruženja u kojem sustav radi, osobito zbog njihove interakcije s fizičkim osobama ili drugim sustavima. U tom se pogledu moraju poduzimati tehničke i organizacijske mjere

Prijava ozbiljnih incidenata (članak 73. Uredbe o umjetnoj inteligenciji)

- dobavljači visokorizičnih UI sustava stavljenih na tržište Unije dužni su prijaviti svaki ozbiljan incident tijelima za nadzor tržišta država članica u kojima je došlo do tog incidenta.

Koje su obveze subjekata koji uvode visokorizične sustave umjetne inteligencije- *AI deployeri* (članak 26. Uredbe o umjetnoj inteligenciji)?

Korištenje i nadzor

- Osigurati korištenje UI sustava u skladu s uputama proizvođača.
- Dodijeliti ljudski nadzor osobama s odgovarajućim znanjem i ovlastima.
- Očuvati autonomiju u organizaciji nadzora u skladu s internim pravilima i zakonom.

Podaci i praćenje rada sustava

- Osigurati da su ulazni podaci relevantni i reprezentativni za svrhu UI sustava.
- Aktivno pratiti rad UI sustava i izvijestiti proizvođača o svim problemima ili rizicima.
- U slučaju ozbiljnog incidenta, odmah obavijestiti dobavljača i nadzorna tijela.

Evidencije i čuvanje zapisa

- Čuvati automatski generirane zapise najmanje 6 mjeseci, ako nije drugačije propisano.
- Financijske institucije čuvaju evidencije u skladu s pravilima o financijskim uslugama.

Obavještavanje radnika

- Poslodavci moraju pravovremeno obavijestiti radnike i njihove predstavnike o primjeni visokorizičnih UI sustava na radnom mjestu.

Obveze javnih tijela

- Javna tijela moraju osigurati registraciju UI sustava u EU bazi podataka.
- Ako sustav nije registriran, ne smije se koristiti, a potrebno je obavijestiti dobavljača.

Procjena učinka na zaštitu podataka

- Ako je primjenjivo, koristiti informacije iz UI sustava za provedbu procjene učinka na zaštitu podataka (čl. 35. Opće uredbe o zaštiti podataka ili čl. 27. Direktive 2016/680).

Posebne obveze – biometrijska identifikacija

Subjekt koji uvodi visokorizični UI sustav za naknadnu daljinsku biometrijsku identifikaciju od pravosudnog ili upravnog tijela čija je odluka obvezujuća i podliježe sudskom preispitivanju traži, *ex ante* ili bez nepotrebne odgode, a najkasnije u roku od 48 sati odobrenje za upotrebu tog sustava osim ako se on upotrebljava za početnu identifikaciju potencijalnog osumnjičenika na temelju objektivnih i provjerljivih činjenica izravno povezanih s kaznenim djelom. Svaka upotreba mora biti ograničena na ono što je nužno za istragu određenog kaznenog djela.

Ako je odobrenje zatraženo na temelju prvog podstavka odbijeno, upotreba sustava za naknadnu daljinsku biometrijsku identifikaciju povezanog s tim zatraženim odobrenjem smjesta se zaustavlja te se brišu osobni podaci povezani s upotrebom visokorizičnog UI sustava za koji je zatraženo odobrenje. Sustavi se ne smiju koristiti u ne ciljane svrhe ili za masovni nadzor bez

konkretne veze s kaznenim djelom. Sve uporabe moraju biti dokumentirane i dostupne nadzornim tijelima na zahtjev (bez otkrivanja osjetljivih operativnih podataka). Obvezna su godišnja izvješća nadzornim tijelima o korištenju takvih sustava.

Informiranje i transparentnost

Ako UI sustav donosi odluke koje utječu na fizičke osobe, korisnici moraju biti informirani da se koristi visokorizični UI sustav. U kaznenom progonu vrijede posebna pravila sukladno Direktivi (EU) 2016/680.

Suradnja s nadzornim tijelima

Subjekti koji uvode UI sustav dužni su aktivno surađivati s nadležnim tijelima u svim postupcima i mjerama vezanim uz provedbu Uredbe.

Procjena učinka na temeljna prava (članak 27. Opće uredbe o umjetnoj inteligenciji)

Prije uvođenja visokorizičnog UI sustava iz članka 6. stavka 2., uz iznimku visokorizičnih UI sustava namijenjenih za upotrebu u području navedenom u točki 2. Priloga III., subjekti koji uvode sustav koji su javnopravna tijela ili su privatni subjekti koji pružaju javne usluge i subjekti koji uvode visokorizične UI sustave iz točke 5. podtočaka (b) i (c) Priloga III. provode procjenu učinka koji upotreba takvog sustava može imati na temeljna prava.

Više o procjeni učinka na temeljna prava možete saznati na poveznici:

- Metodologija za provedbu procjene učinka na ljudska prava.

Kako prepoznati visokorizični sustav umjetne inteligencije (članak 6. Uredbe o umjetnoj inteligenciji)?

Klasifikacija rizika temelji se na namjeni UI sustava, u skladu s postojećim zakonodavstvom EU-a o sigurnosti proizvoda. To znači da klasifikacija ovisi o funkciji koju UI sustav obavlja te o specifičnoj namjeni i načinu korištenja sustava.

UI sustavi mogu biti klasificirani kao visokorizični u dva slučaja:

1. **Ako je UI sustav ugrađen kao sigurnosna komponenta** u proizvode obuhvaćene postojećim zakonodavstvom o proizvodima (Prilog I), ili ako sam predstavlja takav proizvod.
   – Primjer: Sustav umjetne inteligencije za medicinsku dijagnostiku.

2. **Ako je UI sustav namijenjen za visokorizičnu primjenu**, kako je navedeno u Prilogu III Uredbe o umjetnoj inteligenciji.
   – Popis uključuje područja kao što su **obrazovanje, zapošljavanje, provedba zakona ili upravljanje migracijama**.

Europska komisija priprema **smjernice za klasifikaciju visokog rizika**, koje će biti objavljene prije nego što nova pravila stupe na snagu.

Ograničeni broj sustava umjetne inteligencije definiranih u prijedlogu smatra se sustavima visokog rizika, jer potencijalno mogu negativno utjecati na sigurnost ljudi ili na njihova temeljna prava (zaštićena Poveljom Europske unije o temeljnim pravima).

U Prilogu III nalaze se popisi sustava umjetne inteligencije visokog rizika, koji se mogu ažurirati kako bi pratili razvoj novih primjena umjetne inteligencije.

Ovi sustavi uključuju i sigurnosne komponente proizvoda koji su obuhvaćeni sektorskim zakonodavstvom Unije. Uvijek će se smatrati sustavima visokog rizika ako su podvrgnuti ocjeni sukladnosti od strane treće strane prema tom sektorskom zakonodavstvu.

Primjeri takvih sustava umjetne inteligencije visokog rizika uključuju, primjerice, sustave koji procjenjuju je li neka osoba podobna za određeni medicinski tretman, za dobivanje određenog posla ili zajma za kupnju stana. Ostali primjeri uključuju sustave koje koristi policija za profiliranje osoba ili za procjenu rizika od počinjenja kaznenog djela (osim ako su zabranjeni prema članku 5.). Sustavi visokog rizika mogu također biti i oni koji upravljaju robotima, dronovima ili medicinskim uređajima.

Početak primjene članka 6. stavka 1. „Pravila o klasifikaciji UI sustava kao visokorizičnih" i odgovarajućih obveza iz Uredbe počinje od 2. kolovoza 2027.

Kako Uredba o umjetnoj inteligenciji regulira biometrijsku identifikaciju?

Uporaba biometrijske identifikacije u stvarnom vremenu na daljinu u javno dostupnim prostorima (npr. prepoznavanje lica putem nadzornih kamera – CCTV) u svrhe provedbe zakona je zabranjena. Međutim, države članice mogu zakonom predvidjeti iznimke koje dopuštaju takvu uporabu u sljedećim slučajevima:

- Aktivnosti provedbe zakona povezane s 16 jasno određenih vrlo ozbiljnih kaznenih djela;
- Ciljano traganje za žrtvama, nestalim osobama, slučajevima trgovine ljudima i seksualnog iskorištavanja;
- Sprječavanje prijetnje životu ili tjelesnoj sigurnosti osoba, ili reakcija na sadašnju ili predvidivu prijetnju terorističkog napada.

Svaka izvanredna uporaba podliježe prethodnom odobrenju sudske ili neovisne upravne vlasti, čija je odluka obvezujuća. U hitnim slučajevima dopuštenje se može izdati unutar 24 sata; ako se odobrenje odbije, svi prikupljeni podaci i rezultati moraju se izbrisati.

Prije uporabe sustava, potrebno je provesti procjenu učinka na temeljna prava, te o tome obavijestiti nadležno tijelo za nadzor tržišta i tijelo za zaštitu podataka. U hitnim slučajevima sustav se može početi koristiti bez prethodne registracije.

Uporaba UI sustava za naknadnu biometrijsku identifikaciju (identifikaciju osoba na prethodno prikupljenim snimkama) osoba pod istragom također zahtijeva prethodno odobrenje sudske ili neovisne upravne vlasti te obavještavanje nadležnih tijela za zaštitu podataka i nadzor tržišta.

Zašto su potrebna posebna pravila za biometrijsku identifikaciju na daljinu?

Biometrijska identifikacija može imati različite oblike.
Biometrijska autentifikacija i verifikacija (npr. otključavanje mobitela otiskom prsta ili provjera identiteta na graničnim prijelazima putem putovnice) nisu regulirane jer ne predstavljaju značajan rizik za temeljna prava.

S druge strane, biometrijska identifikacija na daljinu, primjerice identifikacija osoba u masi, može značajno utjecati na privatnost u javnim prostorima.

Točnost sustava za prepoznavanje lica ovisi o mnogim čimbenicima: kvaliteti kamere, svjetlu, udaljenosti, bazi podataka, algoritmu te osobinama poput etničke pripadnosti, dobi ili spola osobe. Isto vrijedi i za prepoznavanje hoda, glasa i druge biometrijske sustave.

Iako je sustav s 99 % točnosti dobar u prosjeku, i dalje postoji rizik da se nevinu osobu nepravedno sumnjiči. Čak i stopa pogreške od 0,1 % može imati veliki učinak kada se koristi na velikim populacijama, primjerice na željezničkim kolodvorima.

Kako Uredba o umjetnoj inteligenciji štiti temeljna prava?

Iako EU i države članice već imaju snažne mehanizme zaštite temeljnih prava i zabrane diskriminacije, kompleksnost i netransparentnost nekih UI sustava ("crne kutije") predstavljaju dodatni izazov.

Ljudski orijentiran pristup UI-u znači da primjena mora biti u skladu s pravom o temeljnim pravima. Uvođenjem zahtjeva za transparentnost i odgovornost u razvoj visokorizičnih UI sustava te jačanjem provedbenih kapaciteta, osigurava se da su ti sustavi od početka razvijani u skladu sa zakonom.

Ako dođe do kršenja prava, nadležna tijela će imati pristup potrebnim informacijama za provođenje istrage o usklađenosti UI sustava s pravom EU-a.

Također, Uredba o umjetnoj inteligencije propisuje obvezu provedbe procjene učinka na temeljna prava za određene korisnike visokorizičnih UI sustava.

Mogu li sustavi umjetne inteligencije biti rodno i rasno pristrani?

Važno je naglasiti da UI sustavi ne stvaraju pristranost, već mogu, ako su pravilno dizajnirani i korišteni, smanjiti pristranosti i diskriminaciju. Navedeno može dovesti do pravednijih i nediskriminatornih odluka, primjerice u zapošljavanju.

Sustavi moraju biti tehnički robusni, kako ne bi proizvodili pristrane rezultate (npr. lažno pozitivne/negativne) koji neproporcionalno pogađaju marginalizirane skupine po osnovama kao što su rasa, spol, dob i druge zaštićene osobine.

Također moraju biti obučavani i testirani na reprezentativnim skupovima podataka kako bi se smanjio rizik od nepravednih pristranosti i omogućilo njihovo prepoznavanje i ispravljanje.

Sustavi moraju biti sljedivi i podložni reviziji, s odgovarajućom dokumentacijom – uključujući podatke korištene za treniranje algoritama, a što je ključno za naknadne istrage.

Sustavi moraju biti redovito nadzirani i ažurirani, a identificirani rizici pravovremeno otklonjeni, i prije i nakon stavljanja na tržište.

Što su regulatorni pješčanici?

Sukladno članku 57. Uredbe o umjetnoj inteligenciji, države članice osiguravaju da njihova nadležna tijela uspostave barem jedno regulatorno izolirano okruženje za umjetnu inteligenciju na nacionalnoj razini, koje mora biti operativno do 2. kolovoza 2026. To izolirano okruženje može se uspostaviti i zajedno s nadležnim tijelima drugih država članica.

Regulatorni pješčanik za umjetnu inteligenciju (UI) predstavlja kontrolirano okruženje u kojem organizacije mogu testirati i razvijati UI sustave uz blisku suradnju s nadležnim regulatornim tijelima. Cilj je osigurati da inovacije budu u skladu s postojećim zakonodavstvom, posebno u pogledu zaštite podataka i etičkih standarda.

Norveška agencija za zaštitu osobnih podataka, *Datatilsynet*, uspostavila je regulatorni pješčanik za UI kako bi potaknula inovacije koje poštuju privatnost. U okviru „pješčanika“ organizacijama različitih veličina i sektora pružaju se smjernice za usklađivanje uz uvjet potpune transparentnosti tijekom evaluacije projekata. Cilj je promicati razvoj inovativnih rješenja koja su etična i odgovorna s aspekta zaštite podataka.

Prema Uredbi o umjetnoj inteligenciji, države članice potiču se na uspostavu regulatornih pješčanika kako bi omogućile testiranje UI sustava u stvarnim uvjetima, uz osiguranje usklađenosti s propisima. Ovi pješčanici omogućuju:

- Testiranje inovacija: Razvoj i evaluaciju UI sustava prije njihove pune implementacije na tržištu.
- Suradnju s regulatorima: Blisku suradnju s nadležnim tijelima kako bi se osigurala usklađenost s pravnim i etičkim standardima.
- Identifikaciju rizika: Prepoznavanje i ublažavanje potencijalnih rizika povezanih s UI tehnologijama.