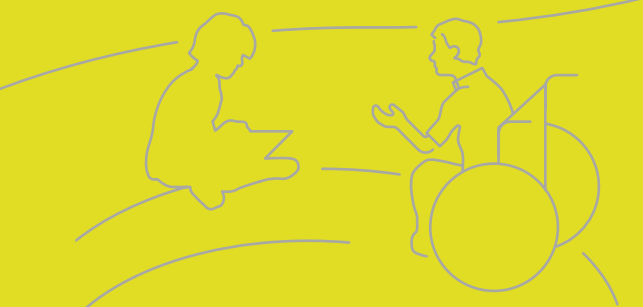


UNESCO  
Priporočilo o

# Etiki umetne inteligence: Oblikovanje prihodnosti naše družbe



**unesco**

Nacionalne komisije  
Nemčije, Nizozemske  
in Slovenije



UNESCO

Priporočilo o etiki  
umetne inteligence:

Oblikovanje prihodnosti  
naše družbe

# Na kratko o UNESCO Priporočilu

UNESCO Priporočilo o etiki umetne inteligence je mednarodno dogovorjen pravni dokument, ki zagotavlja, da vsaka organizacija, podjetje ali posameznik, ki razvija ali uvaja umetno inteligenco, to počne na etičen način in v skladu s človekovimi pravicami.

Etika na področju umetne inteligence pomeni zagotoviti, da naša interakcija s sistemi umetne inteligence ni škodljiva, temveč prispeva k miru, človekovemu dostojanstvu, trajnostnemu razvoju in varnosti.

To bo mogoče doseči le, če bodo otroci in odrasli po vsem svetu imeli priložnost doseči ustrezno raven pismenosti na področju tako tehnološke kot človeške razsežnosti umetne inteligence.

Priporočilo je zato smernica za vse vlade (nacionalne in lokalne) pri oblikovanju zakonov in strategij na področju umetne inteligence.

UNESCO redno izdaja priporočila. To so normativni instrumenti, pri katerih vsako besedo posebej določi vseh 193 držav članic UNESCO.

Po sprejetju Priporočila morajo države članice svetovni skupnosti poročati, na kakšen način so poskrbele za izvajanje dogovorjene politike v svoji državi.

---

## Zakaj potrebujemo globalen dogovor o umetni inteligenci, osredinjeni na človeka?

→ Umetna inteligenca je postala del našega vsakdana. Ima ogromen potencial za družbeno dobro in lahko prispeva k doseganju ciljev trajnostnega razvoja. Z umetno inteligenco je na primer mogoče predvideti pridelek in s tem zagotoviti prehransko varnost. Kljub temu pa uporaba umetne inteligence prinaša tudi tveganja. V številnih primerih so se v zvezi s sistemi umetne inteligence pojavili etični pomisleki – ker lahko prispevajo k povečevanju razlik med spoloma, škodljivo vplivajo na zasebnost ljudi ali omogočajo manipulacijo medijskih vsebin. Te težave se pojavljajo v vseh državah, hkrati pa se skoraj vsi sistemi umetne inteligence uporabljajo na svetovni ravni, zato potrebujemo čezmejni pogled na etični razvoj in uporabo umetne inteligence. Novembra 2021 je 193 držav članic UNESCO sprejelo prvi globalen dogovor o umetni inteligenci, osredinjeni na človeka, in sicer Priporočilo o etiki umetne inteligence.

---

## Kateri so temeljni cilji Priporočila?

→ Priporočilo je zasnovano kot okvir za sisteme umetne inteligence, ki bodo delovali v dobro človeštva in okolja ter za preprečevanje škode. Gre za globalno sprejet pravni dokument, v katerem so opredeljene vrednote in načela ter način njihovega uresničevanja v praksi. Pri tem državam članicam UNESCO predlaga konkretne naloge za ukrepanje na enajstih področjih, vključno z zdravjem, okoljem in enakostjo spolov. Cilj Priporočila je zaščititi in spodbujati človekove pravice in temeljne svoboščine ter človekovo dostojanstvo in enakost. Etična vprašanja v zvezi s sistemi umetne inteligence so pomembna v vseh fazah življenjskega cikla takih sistemov: od raziskav, zasnove in razvoja do uvajanja in uporabe, vključno z vzdrževanjem, delovanjem, spremljanjem in vrednotenjem ter prenehanjem.

→ Priporočilo je soglasno potrdilo vseh 193 držav članic UNESCO. Poseben poudarek daje potrebam in prispevku držav z nizkimi ali srednje visokimi prihodki za etični razvoj in uporabo umetne inteligence. Obenem je razvoj humanocentrične umetne inteligence opredeljen kot naloga celotne družbe, ki daleč presega zgolj regulacijo tehnologije. V Priporočilu je poudarjeno, da morata biti spoštovanje in spodbujanje človekovih pravic podlaga za vsako ureditev o umetni inteligenci.

→ Priporočilo celovito obravnava vsa področja človeškega življenja, na katera vpliva umetna inteligenca. Poseben poudarek je na tistih, ki so bila doslej v razpravah o umetni inteligenci zapostavljena, **na primer ...**

...  
**različne stopnje  
razvitosti držav,**

...  
svetovni jug,

...  
marginalizirane/  
ranljive  
skupine,

...  
pristranskost na  
podlagi spola in  
diskriminacija,

...  
trajnostni  
razvoj.



## Kaj je »umetna inteligenca«?

→ Sistemi umetne inteligence znajo obdelovati podatke in informacije na način, ki je podoben človeškim miselnim procesom ali se zdi, da jih posnema, vključno z vidiki sklepanja, učenja, zaznavanja, predvidevanja, načrtovanja ali nadzora.

Deloma delujejo na podlagi t. i. strojnega učenja (angl. machine learning) in strojnega sklepanja (angl. machine reasoning).



Za zagotovitev etičnega razvoja in uporabe umetne inteligence so v Priporočilu določene zahteve na enajstih področjih. V nadaljevanju je izbor določb, ki prikazujejo področje uporabe Priporočila:

- 1 Zakoni in strategije upravljanja podatkov morajo varovati pravico do zasebnosti, obravnavati vprašanja, kot je nadzor (na podlagi umetne inteligence), in zagotoviti, da posamezniki ohranijo pravice nad svojimi podatki.
- 2 Za reševanje nujnih razvojnih izzivov in zmanjšanje digitalnega razkoraka med svetovnim jugom in svetovnim severom je treba okrepiti mednarodno sodelovanje na področju umetne inteligence.
- 3 Skozi celoten življenjski cikel sistemov umetne inteligence je treba izvajati ocene njihovega etičnega učinka, da se prepreči škoda za ljudi in zagotovi spodbujanje človekovih pravic z umetno inteligenco.
- 4 Postavljeni morajo biti mehanizmi upravljanja, npr. imenovanje pooblaščenec za etiko umetne inteligence, s katerimi bo mogoče raziskati morebitno škodo, ki jo povzročijo sistemi umetne inteligence, sprejeti postopke za uveljavljanje pravic ter zagotoviti pravna sredstva in popravne ukrepe.
- 5 Uvesti je treba spodbude za razvoj z umetno inteligenco podprtih aplikacij za spremljanje, varovanje in obnovo okolja in ekosistemov.
- 6 Umetno inteligenco je treba razviti in uporabljati tako, da bo čim bolj povečala koristi za raznolikost in vključenost, vključno s kulturno raznolikostjo, zagotavljala nediskriminacijo, spodbujala svobodo izražanja in dosegla enakost spolov.

- 7 Politike morajo stremeti k zagotavljanju ustreznega izobraževanja javnosti glede pismenosti na področju umetne inteligence, da bi okrepili vlogo ljudi in zmanjšali digitalni razkorak.
- 8 Sistemi umetne inteligence morajo izboljšati dostop do informacij in znanja.
- 9 Vlade in družbe zasebnega sektorja morajo zagotoviti pravičen prehod za zaposlene, katerih delovna mesta so ogrožena zaradi novih tehnologij in sistemov umetne inteligence.
- 10 Pripraviti je treba smernice za interakcije med ljudmi in roboti ter njihov vpliv na odnose med ljudmi, zlasti ob upoštevanju duševnega in telesnega zdravja ljudi.



# Pet ključnih področij, na katerih Priporočilo poziva k ukrepanju

Priporočilo postavlja jasen etični okvir, znotraj katerega lahko nacionalne vlade razvijejo konkretne politike za izvajanje. V tem pomembnem procesu prenosa v nacionalno »trdo« in »mehko« pravo igrata civilna družba in zasebni sektor ključno vlogo. V naslednjih odstavkih je predstavljenih pet področij in konkretni mandati za politično ukrepanje, da bi zagotovili etičen razvoj in uporabo umetne inteligence.



# Zdravje

V zdravstvu so potencialna tveganja in priložnosti umetne inteligence bolj opazna kot na drugih področjih. Umetna inteligenca že zdaj prinaša revolucijo pri zgodnjem odkrivanju bolezni, odpira nove možnosti zdravljenja in je lahko priložnost za reševanje vse večjega pomanjkanja osebja na področju (zdravstvene) oskrbe. Obenem pa se sektor sooča tudi s pomembnimi etičnimi vprašanji, npr. z varstvom osebnih zdravstvenih podatkov in vmesniki med možgani in računalniki. Ta problematika bo neposredno vplivala na administrativno osebje v zdravstvenih ustanovah, bolnike in zdravnike.

Vlade in deležnike na področju zdravstva pozivamo, naj se posvetijo tem vprašanjem:

- razvoj smernic za interakcije med ljudmi in roboti ter njihov vpliv na medčloveške odnose, s posebnim poudarkom na duševnem in telesnem zdravju ljudi, tako da bodo te interakcije med ljudmi in roboti v skladu s človekovimi pravicami, temeljnimi svoboščinami, spodbujanjem raznolikosti in zaščito ranljivih oseb; zlasti pri uporabi z umetno inteligenco podprtih sistemov za nevrotehnologije in vmesnike med možgani in računalniki;
- uporaba sistemov umetne inteligence za varovanje zdravja ljudi s spremljanjem izbruhov bolezni in izmenjavo podatkov na svetovni ravni;
- spodbujanje skupnega raziskovanja učinkov interakcije ljudi s sistemi umetne inteligence na dolgi rok, s posebnim poudarkom na morebitnem psihološkem in kognitivnem vplivu teh sistemov na otroke in mlade.

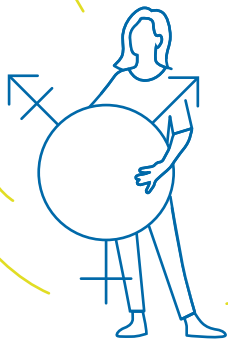
# Kultura

Umetna inteligenca močno vpliva na to, kako »proizvajamo« kulturo, kako in koliko potrošniki plačujejo za kulturne dobrine in kako lahko ustvarjalci živijo od svojega dela. Algoritmi že zdaj posegajo v naše »konzumiranje« kulture in odločajo o tem, ter celo vplivajo npr. na naš izbor glasbe.

Nekatere ključne odgovornosti držav članic UNESCO so:

- sodelovanje s tehnološkimi podjetji in drugimi deležniki za spodbujanje raznolike ponudbe kulturnega izražanja in pluralnega dostopa do njega, ob hkratnem zagotavljanju prepoznavnosti in možnosti odkrivanja lokalnih vsebin;
- spodbujanje novih raziskav na presečišču med umetno inteligenco in intelektualno lastnino, da bi ugotovili, ali in kako bi bilo treba dela, ki jih je ustvarila umetna inteligenca oz. so nastala z njeno pomočjo, zaščititi s pravicami intelektualne lastnine;
- obravnava vpliva sistemov umetne inteligence – zlasti aplikacij za obdelavo naravnega jezika, kot so avtomatsko prevajanje in glasovni pomočniki – na nianse človeškega jezika in kulturnega izražanja.





## ● Spol

Uporaba sistemov umetne inteligence lahko privede do pristranskosti ali jo okrepi. Kadar organizacije uporabljajo prevajalsko programsko opremo, lahko iz spolno nevtralnega izvornika (»the doctor«) nastane spolno zaznamovan prevod (»doktor«). Prav tako je velika verjetnost, da bodo podjetja v panogi, v kateri prevladujejo moški, pri zaposlovanju novih kandidatov na podlagi podatkovnih nizov, ki vsebujejo do sedaj pridobljene življenjepise, reproducirala pristranskost do kandidatk. Eno od pomembnih načel Priporočila je čim bolj zmanjšati in preprečiti krepitev ali nadaljevanje diskriminatornih ali pristranskih aplikacij in rezultatov v celotnem življenjskem ciklu sistema umetne inteligence.

Oblikovalci politik, delodajalci na področju raziskav in razvoja umetne inteligence ter drugi akterji na področju umetne inteligence si morajo prizadevati za:

- vključitev akcijskega načrta za enakost spolov v nacionalne digitalne politike;
- povečanje naložb na ključnih področjih, kot je izobraževanje delavcev z namenom, na primer, spodbujanja poklicnega razvoja žensk v naravoslovju, tehnologiji, inženirstvu in matematiki ter IKT;
- zagotavljanje, da se spolni stereotipi in različne vrste diskriminatorne pristranskosti ne prenesejo v sisteme umetne inteligence (npr. s podporo pri oblikovanju raznolikih, odprtih in zaupanja vrednih podatkovnih nizov, ki bodo predstavljali »zlati standard«).

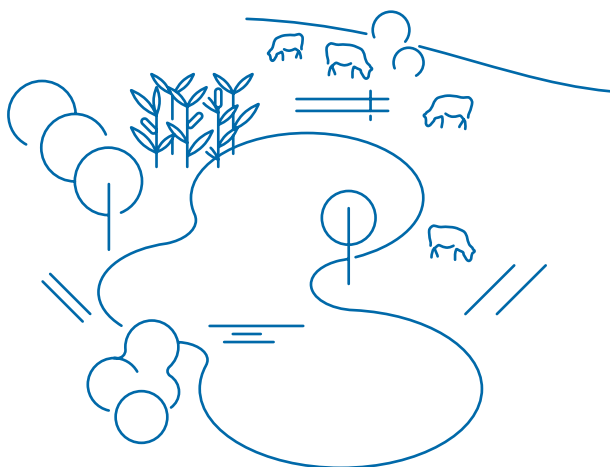


## ● Izobraževanje in raziskave

Šole in zunajšolska učna okolja so bistvena za večanje ozaveščenosti o etiki umetne inteligence in del prizadevanj za premostitev digitalnega razkoraka. V Priporočilu je izobraževanje izpostavljeno kot pomembno sredstvo za razvijanje pismenosti na področju umetne inteligence pri otrocih in odraslih po vsem svetu, ki mora zajemati tako tehnološko kot človeško razsežnost umetne inteligence. Po drugi strani pa izvajanje e-učenja v šolah prinaša tudi etične dileme. Na primer uporaba osebnih podatkov dijakov in študentov mora potekati v skladu z ustreznimi standardi za varstvo osebnih podatkov, pri čemer je treba skrbno preučiti, kdaj in kako uporabljati z umetno inteligenco podprte aplikacije.

Države članice, pa tudi delodajalce v izobraževanju, učitelje in nevladne organizacije pozivamo, naj:

- spodbujajo splošne programe ozaveščanja o razvoju umetne inteligence in vplivu sistemov umetne inteligence na človekove pravice otrok;
- podpirajo medsebojno sodelovanje na področju izobraževanja o tehničnih znanjih in spretnostih in etiki umetne inteligence ter spodbujajo interdisciplinarne raziskave umetne inteligence z vključevanjem tudi drugih področij, ne zgolj naravoslovja, tehnologije, inženirstva in matematike;
- podpirajo znanstveno-politične smernice za politiko glede umetne inteligence in pomagajo ozaveščati o etiki umetne inteligence.



## ● Okolje in ekosistemi

Umetna inteligenca že pomaga varovati okolje. Uporablja se npr. za spremljanje in izboljševanje biotske raznovrstnosti z analizo posnetkov iz brezpilotnih letalnikov in satelitov. Lahko pa sistemi umetne inteligence tudi negativno vplivajo na okolje, npr. zaradi velikih potreb po energiji, ki jo umetna inteligenca potrebuje za svojo računsko moč.

Priporočilo med drugim poziva k ukrepanju v zvezi s temi vprašanji:

- ocenjevanje neposrednega in posrednega vpliva na okolje skozi celoten življenjski cikel sistema umetne inteligence ter skladnost z okoljskimi predpisi;
- zagotavljanje razvoja in sprejemanja na pravicah utemeljenih in etičnih aplikacij, podprtih z umetno inteligenco, za odpornost proti tveganjem nesreč, spremljanje, zaščito in obnovo okolja in ekosistemov ter ohranitev našega planeta;
- prepoved sistemov umetne inteligence, pri katerih obstaja velika verjetnost nesorazmerno visokih negativnih vplivov na okolje.



# Čas za podporo etičnega in k človeku usmerjenega razvoja umetne inteligence

Cilj Priporočila je zagotoviti, da se umetna inteligenca razvija v korist družbe. Da bi razvoj umetne inteligence, usmerjen k človeku, postal in tudi ostal norma, bodo morali prispevati vsi deležniki: ne le oblikovalci politik, temveč tudi organizacije civilne družbe, zasebna podjetja in splošna javnost. Kljub pozivom h konkretnim ukrepom Priporočilo izpostavlja tudi nujnost skupnih prizadevanj. To priložnost za soustvarjanje sistemov, ki zagotavljajo etični in k človeku usmerjen razvoj in uporabo umetne inteligence, bi morali izkoristiti vsi akterji na področju umetne inteligence, ne samo tisti, ki delujejo na področju zdravstva, kulture ali izobraževanja. Čaka nas še kar nekaj dela!

# Kolofon

## Izdajatelj


---

Mednarodni raziskovalni center za umetno inteligenco pod okriljem UNESCO (IRCAI)  
Jožef Stefan Institute  
Jamova cesta 39  
1000 Ljubljana, Slovenija

Slovenska nacionalna komisija za UNESCO  
Kotnikova 38  
1000 Ljubljana, Slovenija

## Kontakt

---

Mednarodni raziskovalni center za umetno inteligenco pod okriljem UNESCO (IRCAI)  
ircai@info.org  
 [www.ircai.org](http://www.ircai.org)

Slovenska nacionalna komisija za UNESCO  
T +386 (0) 1-4784799  
unesco.mizs@gov.si  
 <https://www.gov.si/drzavni-organi/ministrstva/ministrstvo-za-visoko-solstvo-znanost-in-inovacije/o-ministrstvu/urad-za-unesco/>

## Uredniki

---

Anne Diessner, Jeannine Hausmann  
(odgovorna), Maximilian Müngersdorff,  
Davor Orlic, Jon Verriet


## Stanje

---

januar 2024

## Avtorske pravice

---

Besedila te publikacije so zaščiteni z licenco Creative Commons Priznanje avtorstva – Deljenje pod enakimi pogoji 4.0 Mednarodna (CC BYSA 4.0).  [www.creativecommons.org/licenses/by-sa/4.0/deed.sl](http://www.creativecommons.org/licenses/by-sa/4.0/deed.sl)

## Oblikovanje in postavitve

---

IRCAI





