

Etični vidiki rabe generativne umetne inteligence v izobraževanju

Ernest Ženko¹

Povzetek

Prispevek se posveča etičnim vidikom uporabe generativne umetne inteligence (Gen-UI) v izobraževanju, ki postaja pomembno področje sodobnega pedagoškega razmisleka. Avtor ugotavlja, da se je uvedba Gen-UI v šolsko in univerzitetno okolje zgodila kot spontani prihod tehnologije, brez načrtovanih postopkov in okvirov vrednotenja, kar sproža številna vprašanja o avtentičnosti, pravičnosti, zasebnosti, odgovornosti in trajnosti. V prispevku so predstavljeni ključni etični okviri UNESCO, OECD in Evropske komisije, ki skupaj oblikujejo jedro globalnega pristopa, ki se usmerja k zaupanja vredni umetni inteligenci. Poudarek je na tem, da etična uporaba Gen-UI ne pomeni omejevanja tehnologije, pač pa njeno umestitev v širši humanistični in pedagoški okvir, kjer tehnologija podpira človekovo dostojanstvo, razvoj avtonomije in pravičnost. Posebej so izpostavljena temeljna načela odgovorne rabe, kot so človeški nadzor, preglednost, odgovornost, pravičnost, varnost in trajnost. V sklepu avtor opozarja, da etika v kontekstu umetne inteligence presega tehnična vprašanja ter predstavlja prizadevanje za ohranjanje človeških vrednot v digitalni dobi: iskanje ravnotežja med tehnološkim determinizmom in etično avtonomijo.

Ključne besede: generativna umetna inteligenca (Gen-UI), etika v izobraževanju, odgovorna raba tehnologije, zaupanja vredna umetna inteligenca, mednarodni etični okviri, človekovo dostojanstvo in avtonomija

Ethical Considerations of the Use of Generative Artificial Intelligence in Education

Abstract

This article focuses on the ethical aspects of using generative artificial intelligence (Gen-AI) in education, which is becoming an important area of contemporary pedagogical reflection. The author notes that the introduction of Gen-AI into school and university environments has occurred as a spontaneous arrival of technology, without planned

1 Fakulteta za Management, Univerza na Primorskem, ernest.zenko@upr.si

procedures and evaluation frameworks, raising numerous questions about authenticity, fairness, privacy, responsibility, and sustainability. The article presents the key ethical frameworks of UNESCO, the OECD, and the European Commission, which together form the core of a global approach aimed at trustworthy artificial intelligence. The emphasis is on the fact that the ethical use of Gen-AI does not mean restricting technology, but rather placing it in a broader humanistic and pedagogical framework where technology supports human dignity, the development of autonomy, and fairness. Particular emphasis is placed on the fundamental principles of responsible use, such as human oversight, transparency, accountability, fairness, security, and sustainability. In conclusion, the author points out that ethics in the context of artificial intelligence goes beyond technical issues and represents an effort to preserve human values in the digital age: finding a balance between technological determinism and ethical autonomy.

Keywords: generative artificial intelligence (Gen-AI); ethics in education; responsible use of technology; trustworthy artificial intelligence; international ethical frameworks; human dignity and autonomy

Uvod: neizogibnost etike

Tehnologije, ki jih poznamo pod skupnim imenom »umetna inteligenca« (UI), niso nikakršna novost, saj lahko njihov razvoj spremljamo že vsaj sedem desetletij; toliko časa je namreč že preteklo, odkar je John McCarthy vpeljal ta izraz (Nilsson, 2013). Njihov vpliv na področju izobraževanja je bil v tem času sicer prisoten, vendar pa ne tako kot v zadnjih nekaj letih, odkar so se v širši javnosti pojavila orodja generativne umetne inteligence (Gen-UI). Ta so namreč pričela spreminjati dogajanje na področju izobraževanja veliko hitreje, kot se je temu uspelo prilagajati ustanovam, učiteljem in učnim načrtom. Orodja, kot so ChatGPT, Gemini, Copilot, DeepSeek ipd., po vsega nekaj letih prisotnosti in rabe niso več eksperimentalna tehnološka novost, pač pa so vsakdanji del učnega procesa: sodelujejo pri pripravi učnega gradiva, preverjanju znanja, pisanju besedil in oblikovanju različnih projektov.

Reich in Dukes (2024) opozarjata na razliko med načrtovanimi (ang. *adopted*) in spontanimi (ang. *arrival*) tehnologijami, ravno razlikovanje med njimi pa po njunem mnenju omogoča boljše razumeti tudi posledice uvajanja različnih tehnologij v izobraževanje. Prve v šole uvajajo postopno – z načrtovanjem, oceno tveganj, pilotnimi projekti in usposabljanjem. Spontane tehnologije, kot je gen-UI, pa vstopijo nenadno in brez uradne odločitve, saj jih v izobraževalni prostor večinoma uvedejo uporabniki. V primeru Gen-UI je bilo to možno zaradi predhodno razširjene digitalne infrastrukture in hitre globalne rabe orodij, učitelji posledično niso bili pripravljeni niti na njen obstoj niti na posledice rabe tehnologije, ki je v šole vstopila brez jasnih vrednostnih in etičnih okvirov.

Zaradi tega se odpirajo številna vprašanja, ki presegajo zgolj tehnično razumevanje tehnologije: Kako zagotoviti avtentičnost učenčevega (in učiteljevega) dela? Kako ohraniti pravičnost in zagotoviti varnost podatkov? Kako preprečiti odvisnost od avtomatiziranih orodij in spodbuditi refleksijo o njihovem vplivu na mišljenje in ustvarjalnost? Etika pri rabi Gen-UI tako ni vezana zgolj na vprašanje, »kaj je dovoljeno?«, saj je v tem kontekstu zastavljeno preozko – tudi prakse, ki s pravnega vidika niso eksplicitno prepovedane, lahko predstavljajo etična tveganja. Tako sta boljši vprašanja – »kaj je prav?« in »kaj je smiselno?«, in sicer v pedagoškem procesu, pa tudi z vidika človeškega pogleda.

Zato uvajanje Gen-UI v izobraževanje zahteva novo vrsto pismenosti – etično pismenost, ki dopolnjuje digitalno in informacijsko pismenost (Ženko, 2025). Ta pomeni sposobnost presojanja učinkov tehnologij na ljudi, odnose in vrednote, ne nazadnje pa tudi na planet in družbo. Učitelji se v dobi Gen-UI ne znajdejo več zgolj v vlogi posrednikov znanja, temveč morajo postati tudi etični moderatorji, kar pa pomeni, da morajo imeti za to novo nalogo tudi ustrezna znanja in kompetence s področja etike.

Pri tem je nujno izhajati iz prepričanja, da etika ni zgolj zunanji dodatek tehnologiji, temveč predvsem njen notranji kompas. Cilj ni omejevanje uporabe umetne inteligence, pač pa njeno smiselno vključevanje v kontekst človekovega dostojanstva, odgovornosti in pravičnosti. Zato bomo v nadaljevanju predstavili nekatere mednarodne etične okvire, izpostavili glavne dileme, ki spremljajo rabo Gen-UI, ter nakazali poti k vzpostavitvi odgovorne rabe v praksi.

Etični okviri

Na nekaterih področjih, kot je zdravstvo, velja, da ne zadostuje individualna etična presoja, prav tako ne presoja posamične ustanove (Ženko, 2024). Tako tudi za etično rabo Gen-UI v izobraževanju ne zadostuje individualna presoja učiteljev ali šol. Raba te tehnologije v procesih učenja in poučevanja odpira vprašanja, ki presegajo osebne vrednote, pedagoško presojo in vsakodnevno prakso, saj vključuje zapletene tehnološke procese, ogromne količine podatkov in algoritme z netransparentnim delovanjem ter težko predvidljivimi posledicami. Posamezniki praviloma nimajo ustreznega vpogleda v tehnične, pravne in družbene vidike teh sistemov, zato velja, da osebna presoja – četudi dobronamerna – ne zadošča za zagotavljanje pravičnosti, varnosti in odgovornosti.

Zato se vprašanja etične presoje nanašajo na ravni sistemskih etičnih okvirov, ki povezujejo vrednote, zakonodajo in prakso ter omogočajo usklajeno, pregledno in odgovorno rabo Gen-UI v izobraževanju. Med najpomembnejše in najvplivnejše sodijo dokumenti UNESCO, OECD in Evropske komisije, ki skupaj tvorijo jedro

globalnega pristopa k »zaupanju vredni umetni inteligenci«. Ponujajo namreč skupen jezik, standarde in mehanizme, ki omogočajo, da posamezne odločitve učiteljev in institucij postanejo del širše etične in pravne zavezanosti. V nadaljevanju so predstavljene ključne značilnosti treh dokumentov, ki oblikujejo normativno in praktično osnovo za odgovorno vključevanje Gen-UI v izobraževanje.

UNESCO: človekovo dostojanstvo in odgovorna raba

UNESCO že več desetletij razvija globalne smernice za etično uporabo tehnologij v izobraževanju, pri čemer izhaja iz humanistične vizije, da mora vsaka oblika tehnološkega napredka služiti človeku in njegovi avtonomiji. V dokumentih *AI and Education: Guidance for Policy-Makers* (2021) in *Guidance for Generative AI in Education and Research* (2023) UNESCO poudarja, da mora biti umetna inteligenca – še posebej generativna – osredinjena na človeka ter usmerjena v podporo razvoju moralnega, ustvarjalnega in avtonomnega posameznika.

Ključna načela tega pristopa so:

- človekovo dostojanstvo kot temeljni kriterij vsake odločitve o uporabi umetne inteligence;
- pravičnost in vključevanje z namenom zmanjševanja digitalnega razkoraka in preprečevanja nove oblike »digitalne revščine«;
- transparentnost in sledljivost, ki zahtevata podatke o tem, kdaj in kako se uporablja UI, ter omogočata preverljivost rezultatov;
- odgovornost in preglednost, kar pomeni, da morajo biti sistemi umetne inteligence podvrženi neodvisnemu nadzoru in možnosti popravljanja napak.

UNESCO v svojih dokumentih opozarja, da se Gen-UI s sposobnostjo posnemanja višjih kognitivnih procesov, kot sta pisanje in ustvarjanje podob, dotika samega bistva učenja in mišljenja. Zato ne gre le za vprašanje orodij, pač pa tudi za ponoven premislek o tem, zakaj, kaj in kako učimo. Generativna orodja lahko učeče se osvobodijo nižjih kognitivnih nalog, hkrati pa od učiteljev zahtevajo refleksijo o tem, kaj v digitalni dobi pomeni ustvarjalnost, razumevanje in znanje.

Dokument iz leta 2023 opozarja tudi na številna tveganja in dileme, ki spremljajo uporabo Gen-UI v izobraževanju:

- digitalna neenakost in dostop do virov,
- pomanjkanje ustrezne nacionalne regulative in zaščite podatkov,
- učenje modelov na avtorskih vsebinah brez soglasja avtorjev,
- netransparentnost modelov in nezmožnost, da bi pojasnili sprejeto odločitev (problem »črne škatle«),
- homogenizacija kulturnih izrazov in zmanjševanje raznolikosti mnenj ter
- pojav t. i. globokih ponaredkov (ang. »deepfakes«), ki ogrožajo zaupanje v znanje in institucije.

V odgovoru na te izzive UNESCO predlaga naslednje konkretne ukrepe:

- Vzpostavitev na človeka osredinjene regulacije, ki daje prednost zaščiti zasebnosti, etičnosti in varnosti.
- Uvedba mehanizmov preverjanja in potrjevanja primernosti orodij Gen-UI v izobraževanju.
- Razvoj kompetenc učiteljev, raziskovalcev in učencev, ki vključujejo tehnično, etično in družbeno razumevanje UI.
- Spodbujanje večglasja in kulturne raznolikosti, z namenom preprečevanja standardizacije znanja in mišljenja.
- Medsektorsko interdisciplinarno sodelovanje pri spremljanju dolgoročnih učinkov Gen-UI na učenje, poučevanje, raziskovanje in človeško identiteto.

Poseben poudarek je v dokumentu namenjen pismenosti učiteljev s področja UI. UNESCO poudarja, da učitelji niso zgolj uporabniki orodij, pač pa tudi njihovi posredniki in kritični interpreti. Zato pismenost s področja UI (in predvsem Gen-UI) vključuje ne le tehnično razumevanje uporabljenih algoritmov, pač pa tudi sposobnost presoje njihovih etičnih in družbenih posledic. Učitelji morajo razumeti omejitve UI – njeno nezmožnost razumevanja realnega sveta in družbenih razmerij – ter učencem pomagati prepoznati razlike med človeškim mišljenjem in rezultati na statistiki temelječega sklepanja strojev.

Zato UI v teh dokumentih ni pojmovana kot zamenjava za človeško inteligenco, temveč kot povabilo k ponovnemu premisleku o znanju, učenju in človeški ustvarjalnosti. Cilj ni avtomatizacija izobraževanja, pač pa oblikovanje novih, humanistično utemeljenih horizontov, v katerih tehnologija krepi človekovo dostojanstvo, kritično mišljenje in družbeno pravičnost.

OECD: na človeka osredinjena umetna inteligenca

OECD v globalnih smernicah za umetno inteligenco poudarja, da mora biti raba UI in Gen-UI v izobraževanju inovativna, zaupanja vredna ter skladna s človekovimi pravicami in demokratičnimi vrednotami (OECD, 2024). To pomeni, da je v izobraževanju ključno spoštovati človekovo dostojanstvo vsakega učenca in učitelja, zagotavljati pravičnost (denimo preprečevati diskriminacijo v algoritmih) ter uveljavljati transparentnost in razložljivost delovanja UI, ob tem pa vzpostaviti jasno odgovornost za odločitve, ki jih takšni sistemi sprejemajo. OECD ta načela uveljavlja z namenom podpiranja dobrobiti učečih se z vidika tehnologije in ne zmanjševanja njihove avtonomije ali pravic, saj morajo orodja UI delovati v službi človeka. Načelo transparentnosti je posebej pomembno za obvladovanje tveganj t. i. algoritmične netransparentnosti, saj je za ohranitev zaupanja v izobraževalni proces nujno, da so odločitve in priporočila Gen-UI preverljiva in razumljiva.

Na področju regulacije in razvoja etičnih okvirov za uporabo Gen-UI v učnih okoljih se OECD zavzema za prilagodljiv pristop (OECD, 2023). Trenutno večina držav OECD še nima posebnih predpisov za rabo Gen-UI v šolah; namesto tega so bile oblikovane neobvezujoče smernice, odločanje o uporabi teh orodij pa je pogosto prepuščeno lokalnim oblastem, izobraževalnim ustanovam in učiteljem (seveda ob spoštovanju širših pravil o varstvu podatkov in digitalni varnosti). OECD spodbuja oblikovanje smernic, ki se sproti posodablajo, in okvirov, ki temeljijo na dejanskih izkušnjah in vključujejo varovalke za etično rabo Gen-UI. Pri tem priporoča uravnoteženost: dokumenti naj poudarijo potencialne izobraževalne koristi Gen-UI, hkrati pa jasno obravnavajo tveganja, kot so pristranskost algoritmov (za zaščito pravičnosti), vprašanja zasebnosti in varnosti podatkov, ter vzpostavljajo nadzor, ki zagotavlja odgovornost odločevalcev. Namesto togih prepovedi zagovarja prilagodljiva pravila z mehkejšimi regulativnimi ukrepi, kot je sprotno posodabljanje smernic in varovalk, ki omogočajo inovacije ob hkratnem zagotavljanju transparentnosti in odgovornosti pri uporabi UI. Zato OECD priporoča tudi vzpostavitev etičnih komisij in svetovalnih teles na ravni šolskih sistemov z namenom ocenjevanja vpliva rabe UI v izobraževanju.

Pri zagotavljanju varne in smiselne vključitve Gen-UI v šole igrajo ključno vlogo tako učitelji kot nacionalne politike. Učitelji so pri uvajanju sprememb na prvem mestu, OECD pa poudarja, da morajo države vlagati v njihovo sistematično usposabljanje in krepitev digitalne ter etične pismenosti, da bodo pedagogi Gen-UI uporabljali premišljeno in kompetentno (OECD, 2023). Programi usposabljanja naj bi bili oblikovani tako, da učiteljem ne bi zgolj predstavili tehnične zmožnosti novih orodij UI (npr. kako lahko generativni modeli spodbudijo učenje ustvarjalnost ali kritično mišljenje), marveč bi obenem opozarjali na meje njihove uporabe – npr., zakaj ni smiselno, da bi UI v celoti prevzela ocenjevanje znanja. Poudarjeno je tudi, da lahko prepuščanje odločitev o Gen-UI zgolj posameznim učiteljem ali šolam vodi v neenotne prakse in poveča digitalno neenakost med učenci, zaradi česar OECD priporoča sistemske rešitve na ravni izobraževalnih politik. Z jasnimi smernicami, ustrezno podporo in infrastrukturo je treba vsem šolam omogočiti varno ter smiselno uporabo Gen-UI (ne zgolj tistim z najbolj tehnološko podkovanim kadrom), pri čemer se lahko države oprejo tudi na globalna priporočila ter si med seboj delijo primere dobrih praks.

Evropska komisija: zaupanja vredna umetna inteligenca

Evropska komisija pri oblikovanju okvirov za umetno inteligenco v izobraževanju izhaja iz temeljnih človekovih pravic in etičnih načel, med katerimi je človekovo dostojanstvo nedotakljiva vrednota, ter poudarja pravičnost, odgovornost in transparentnost kot ključne pogoje pri uporabi UI (Evropska komisija, 2019). Taka načela se zrcalijo v mednarodnih smernicah, ki zahtevajo, da vsaka odločitev o uporabi UI

spoštuje dostojanstvo posameznika, zagotavlja pravičnost in vključevanje, omogoča transparentnost (sledljivost) ter uvaja mehanizme odgovornosti in neodvisnega nadzora. Evropska komisija s tem prevzema vodilno vlogo pri krepitvi človeku prijazne in zaupanja vredne umetne inteligence v šolah, kjer tehnologija služi avtonomiji posameznika in varstvu njegovih pravic, namesto da bi jih ogrozila.

V *Etičnih smernicah za zaupanja vredno umetno inteligenco* (2019), ki jih je pripravila neodvisna skupina na visoki ravni za umetno inteligenco, je predstavljen okvir za zaupanja vredno umetno inteligenco, ki jo sestavljajo trije elementi:

- zakonitost: zagotavljanje spoštovanja vseh veljavnih zakonov in drugih predpisov;
- etičnost: zagotavljanje spoštovanja etičnih načel in vrednot ter
- robustnost s tehničnega in družbenega vidika: ker lahko sistemi UI povzročijo nenamerno škodo, tudi če se uporabljajo z dobrimi nameni.

Zaupanja vredna UI temelji na štirih etičnih načelih, med katerimi ne sme biti trenj:

- spoštovanje človekove avtonomije;
- preprečevanje škode;
- pravičnost in
- razložljivost.

Da bi lahko zaupanja vredno UI uresničili, dokument predlaga sedem ključnih zahtev, ki jih je treba nenehno ocenjevati ter obravnavati s tehničnimi in netehničnimi metodami v celotnem življenjskem ciklu sistema UI:

- človekovo delovanje in nadzor;
- tehnična robustnost in varnost;
- zasebnost in upravljanje podatkov;
- preglednost;
- raznolikost, nediskriminacija in pravičnost;
- družbena in okoljska blaginja;
- odgovornost.

Vpliv pristopa, ki v ospredje postavlja zaupanja vredno UI, se v šolskem okolju kaže v zahtevi, da morajo sistemi UI krepiti avtonomijo in dobrobit učencev ter učiteljev (ne pa je spodkopavati), varovati otroke in mladostnike pred škodljivimi vplivi in pristranskostjo, spodbujati enake možnosti ter delovati pregledno in razložljivo. Smernice posebej poudarjajo potrebo po človeškem nadzoru (UI naj podpira pedagoške odločitve, ne sme pa samostojno ukazovati) ter inkluzivnosti, saj je treba preprečiti diskriminacijo ali poglobljanje digitalnega razkoraka. Takšen etični kompas v izobraževanju naj bi vzpostavljal zaupanje v UI ter razvijalce in izvajalce usmerjal, da tehnologijo umeščajo v šolski prostor skladno z vrednotami vzgoje in izobraževanja.

Z *Aktom o umetni inteligenci* (2024) je Evropska komisija etične usmeritve nadgradila v pravno zavezujoč sistem razvrščanja sistemov UI glede na tveganje. Takšen pristop odraža prepričanje, da je dopustnost rabe UI odvisna od njenega vpliva na človekove pravice in družbo: najbolj sporni in nedopustni načini uporabe (npr. manipulacija otrok z vedenjskimi algoritmi ali socialno točkovanje posameznikov) so v EU prepovedani, sistemi UI na občutljivih področjih – med katere sodi tudi šolstvo – pa opredeljeni kot visoko tvegani, kar terja strožji nadzor in preverjanje skladnosti z etičnimi standardi. Tudi orodja generativne UI ne smejo delovati brez omejitev, saj zanje veljajo obveznosti transparentnosti ter mehanizmi za preprečevanje nezakonitih ali škodljivih učinkov.

Tako oblikovan evropski okvir zaupanja vredne UI ne pomeni le pravnega nadzora nad to tehnologijo, temveč izraža tudi etično zavezo, da mora biti tehnologija podrejena človeku kot razumnemu in moralnemu bitju. V kontekstu izobraževanja to pomeni, da Gen-UI ne sme nadomestiti kritičnega mišljenja, ustvarjalnosti in avtonomije, temveč jih mora krepiti.

Etične dileme in izzivi v izobraževalni praksi

Gen-UI v izobraževanju prinaša številne potencialne koristi, od pomoči pri pripravi učnega gradiva do zmanjševanja administrativnih bremen, obenem pa odpira nova etična vprašanja. V šolskem prostoru se ta nanašajo predvsem na razmerje med človekom in tehnologijo ter na vpliv algoritmov na procese odločanja, mišljenja, presoje in ustvarjalnosti. V nadaljevanju navajamo nekatere dileme in izzive, ki se pri rabi Gen-UI v tem kontekstu pogosto pojavljajo.

Avtentičnost rezultatov in vprašanje avtorstva

Generativna orodja omogočajo izdelavo besedil, slik, kode ali videov z minimalnim človeškim vložkom, pri čemer odpirajo zelo pomembno vprašanje: Kdo je avtor tako nastalega izdelka? Ali je seminarsko nalogo, diplomsko delo ali celo doktorsko disertacijo, ki so bili izdelani s pomočjo orodja Gen-UI, še vedno mogoče pojmovati kot rezultat dela učenca ali študenta? Izobraževalne ustanove po vsem svetu uvajajo različne prakse – od obveznega razkritja uporabe Gen-UI do prepovedi pri določenih nalogah (in celo do popolne svobode pri rabi teh orodij). Etično ravnanje v tem primeru pomeni transparentnost in dokumentiranje rabe UI kot orodja in ne kot nadomestka miselnega procesa.

Zasebnost in varstvo podatkov

Izobraževalne ustanove pogosto uporabljajo brezplačne komercialne platforme, ki zbirajo in analizirajo uporabniške podatke, saj so večino orodij Gen-UI razvile korporacije, katerih osnovna »skrb« je dobiček. Ker številne takšne platforme niso skladne z GDPR, obstaja tveganje, da se osebni podatki učencev oz. študentov uporabljajo v komercialne ali analitične namene. Tipičen primer slabe prakse je uporaba brezplačnih orodij za preverjanje znanja ali generativnih aplikacij, ki ob nalaganju učenčevih besedil podatke shranijo na zunanje strežnike in jih uporabijo za nadaljnje učenje modelov. Podobno se dogaja pri uporabi učnih aplikacij, ki zahtevajo prijavo z e-poštnim naslovom in tako ustvarjajo digitalne profile otrok. Zato mora uporaba Gen-UI v izobraževanju slediti načelom minimalnega zbiranja podatkov, izrecne in informirane privolitve ter lokalnega ali varno anonimiziranega hranjenja informacij oz. uporabe licenčnih sistemov, sicer lahko pride do kršitev zasebnosti in posledično do zmanjševanja zaupanja v digitalno izobraževanje.

Algoritemska pristranskost in diskriminacija

Veliki jezikovni modeli se učijo na ogromnih količinah podatkov (besedil, slik, glasbe), ki odražajo družbene pristranskosti – spolne, etnične, kulturne in jezikovne. To v kontekstu izobraževanja pomeni, da lahko generativna orodja reproducirajo stereotipe ali neenakosti. Tako npr. moška imena pogosteje povezujejo z voditeljskimi vlogami, ženska pa z »mehkejšimi« poklici, ali slabše razumejo učence, ki ne uporabljajo angleščine kot maternega jezika. Tveganje predstavlja tudi samodejno ustvarjanje učnih gradiv, kjer se pojavljajo kulturno pristranske ali poenostavljene predstave družbenega konteksta. Posledično je nujno redno preverjanje uporabe orodij Gen-UI – ne le tehnično, temveč predvsem vsebinsko, z vidika pravičnosti, zastopanosti in vpliva na učence. Pomembno je tudi, da učitelji ohranijo kritično presojo rezultatov in ne prenašajo odgovornosti na orodja Gen-UI.

Kognitivni dolg in odvisnost

Nekatere nedavne raziskave (Kosmyna idr., 2025) opozarjajo, da lahko stalna uporaba Gen-UI vodi v pojav t. i. kognitivnega dolga (ang. *cognitive debt*), pri katerem uporabniki postopno izgubljajo mentalno angažiranost in sposobnost samostojnega reševanja kompleksnih nalog. Omenjena raziskava je pokazala, da posamezniki, ki pri pisanju esejev uporabljajo ChatGPT kot primarno orodje, hitreje sprejemajo rešitve brez kritičnega razmisleka in se redkeje vračajo k preverjanju dejstev ali strukturiranju argumentov. To v izobraževalnem kontekstu

pomeni, da lahko učenci ali študenti zlahka razvijejo odvisnost od algoritmične podpore in pri tem ne krepijo ključnih spretnosti: analize, presoje, ustvarjalnosti in vztrajnosti pri miselnem naporu. Zato etika uporabe Gen-UI pomeni zahtevo po pedagoški uravnoteženosti med uporabo in refleksijo: UI naj spodbuja miselni proces, ne sme pa ga nadomestiti.

Digitalni razkorak

Čeprav na prvi pogled orodja Gen-UI delujejo enakopravno in so dostopna »vsem«, se v praksi pokaže, da obstajajo znatne razlike, ki se nanašajo na kako-vost dostopa do spleta, zmogljivost računalnika ali telefona, mirno in varno okolje za delo, pa tudi na znanje tujih jezikov ali digitalne pismenosti, potrebne za rabo teh orodij. Raziskave jasno kažejo, da je prav pri uporabi UI v izobraževanju digitalna neenakost izrazitejša kot sicer: infrastruktura, oprema, usposobljenost učiteljev in učencev ter institucionalna podpora so pogoj za uspeh (Lourdu idr., 2025; Abrahamson, 2025). Zato je z vidika etike nujno, da sta pri uvajanju orodij Gen-UI v izobraževalne politike in prakse upoštevani pravičnost in vključenost. Ni pomembno le to, da orodje obstaja, pač pa tudi, da je dostopno, razumljivo in uporabno za vse deležnike. Le takšen pristop je v skladu z načeli o pravičnem, vključujočem in dostopnem izobraževanju (Gonzales, 2024).

Etika ocenjevanja

V času, ko se meje med človeškim in strojno generiranim delom brišejo, mora ocenjevanje znanja vključevati tudi proces in ne samo izdelka. Kot kažejo nedavne študije, integracija Gen-UI v pisanje nalog pomeni, da lahko učenci zgolj »oddajo« izdelek brez poglobljene refleksije o tem, kako so do njega prišli, kar zmanjšuje njihov metakognitivni angažma in sposobnost reševanja problemov (Xia idr., 2024).

Ob ustreznem pristopu lahko raba Gen-UI postane priložnost za učenje o zanesljivosti virov, avtentičnosti, odgovornosti in lastnem procesu razmišljanja, a le, če je ta postopek transparenten in reflektiran. To pomeni, da mora učitelj oz. mentor pojasniti, kdaj se orodja Gen-UI uporabljajo kot orodje za pomoč (npr. generiranje idej, strukture ali prve skice), kdaj pa se zahteva lastna produkcija, kako naj učenec označi orodje in njegovo rabo ter kako reflektira proces dela.

Načela odgovorne rabe generativne umetne inteligence v izobraževanju

Etična raba Gen-UI pri učenju in poučevanju ne pomeni prepovedi, temveč vzpostavitev varovalnih mehanizmov, ki omogočajo, da tehnologija podpira človeške vrednote in cilje izobraževanja. UNESCO, OECD in Evropska komisija poudarjajo,

da morajo biti ta načela vgrajena v prakso – v učne načrte, študijske programe, evalvacijske procese in institucionalne politike. V nadaljevanju so predstavljena nekatera temeljna načela, ki izhajajo iz dokumentov omenjenih institucij ter tvorijo ogrodje odgovorne rabe Gen-UI v izobraževanju.

Načelo človeškega nadzora

Vsak učni proces, ki vključuje UI, mora ohraniti človeški nadzor kot končno instanco odločanja. Učitelj in učenec nista le uporabnika, temveč tudi kritična interpretata nastalih rezultatov. Odgovornosti za pedagoške odločitve ni mogoče prenesti na sistem. To v praksi pomeni, da Gen-UI npr. lahko pomaga pri pripravi učnega gradiva, toda učitelj presoja, ali so ta ustrezna, smiselna in etično sprejemljiva. To načelo preprečuje prenos mišljenja na »zunanj sistem«, v katerem bi algoritmi pričeli določati didaktične strategije brez človeške refleksije.

V tem kontekstu je treba razumeti tudi vlogo učiteljev in učencev oz. študentov pri rabi orodij Gen-UI. Učitelj je prvi »filter« med tehnologijo in učencem, zato njegova naloga ni le poučevanje o rabi Gen-UI, pač pa tudi vzgoja za odgovorno rabo, ki vključuje razpravo o avtorstvu, plagiatstvu in ustvarjalnosti, refleksijo o vplivu orodij Gen-UI na mišljenje in vrednote (Ženko, 2025) ter oblikovanje jasnih pravil rabe Gen-UI pri delu. Dober primer je finski *Vodnik po umetni inteligenci za učitelje*, ki ponuja nabor etičnih scenarijev in vprašanj, s katerimi lahko učitelji spodbudijo in vodijo dialog o orodjih Gen-UI (Faktabaari, 2025).

Vendar pa tudi učenci oz. študenti niso le uporabniki orodij, pač pa morajo biti aktivni soustvarjalci, ki z vključevanjem v razprave o Gen-UI razvijajo kritično pismenost na tem področju ter občutek odgovornosti. Na nekaterih ustanovah (npr. Univerza v Calgaryju) so oblikovali posebna orodja, ki pomagajo študentom razumeti in ovrednotiti njihovo delo z Gen-UI (McDermott in Anselmo, 2024).

Načelo preglednosti oz. transparentnosti

Uporaba Gen-UI v izobraževanju mora biti pregledna in jasno razkrita, tako pri delu učiteljev kot tudi učencev oz. študentov. Pomembno je, da vsi deležniki – učenci, učitelji in širša javnost – vedo, kdaj in kako je bila UI uporabljena pri učenju, ocenjevanju ali pripravi nalog. To še posebej velja takrat, ko učitelji uporabljajo Gen-UI pri oblikovanju navodil in povratnih informacij ali ocenjevanju, pa tudi, ko učenci oz. študenti pri pisanju nalog posegajo po teh orodjih, kar lahko odpira vprašanja goljufanja, avtorstva in avtentičnosti znanja. Preglednost povečuje zaupanje in omogoča kritično presojo rezultatov. Zato bi morale šole v skladu s priporočili UNESCO (2023) vzpostaviti jasno politiko uporabe Gen-UI, ki določa pravila o njeni rabi in ustreznem navajanju.

Načelo odgovornosti

Načelo odgovornosti pomeni, da za etično rabo UI v izobraževanju ni odgovoren le učitelj, temveč celoten ekosistem – od razvijalcev in ponudnikov orodij do ustanov, ki tehnologijo uvajajo in nadzorujejo njeno uporabo. Pri tem je treba jasno določiti, kdo prevzema odgovornost v primeru napak, napačnih informacij ali pristranskih rezultatov, ki jih ustvari umetna inteligenca. Zato OECD (2024) priporoča, da izobraževalne ustanove uvedejo etično presojo učinkov (Ethical Impact Assessments), obenem s presojo varnosti podatkov. Takšen pristop spodbuja sistemsko odgovornost in pomaga preprečevati situacije, v katerih bi se posameznik znašel pred posledicami odločitev, ki jih je oblikoval sistem UI ali institucija, ki ga uporablja.

Načelo pravičnosti

Gen-UI mora pri učenju in poučevanju delovati nepristransko in zagotavljati enak dostop do znanja za vse učence, ne glede na spol, jezik, socialni položaj ali kulturno ozadje. To zahteva preprečevanje algoritmične pristranskosti in skrbno izbiro orodij, ki upoštevajo raznolikost učnega okolja. To v praksi pomeni, da šole dajejo prednost rešitvam, ki temeljijo na odprtokodnih načelih, vključujejo lokalne podatkovne baze ter omogočajo nadzor nad učnimi podatki in njihovo uporabo. Pravičnost v izobraževanju ni le tehnično, temveč tudi družbeno vprašanje: umetna inteligenca naj ne pogloblja že obstoječih razlik, temveč spodbuja vključevanje in zmanjšuje ovire pri učenju. Na to opozarjajo tudi smernice UNESCO (2023) in OECD (2024), ki poudarjajo, da je pravičnost eden temeljnih pogojev za zaupanje v UI v izobraževanju.

Načelo varnosti in zasebnosti

Varstvo podatkov je eno temeljnih področij etične odgovornosti pri uporabi UI v izobraževanju. Tehnologija mora biti zasnovana tako, da varuje zasebnost učencev, učiteljev in institucij ter preprečuje zlorabo osebnih informacij. Kadar je mogoče, naj se orodja uporabljajo na lokalnih ali na zaupanja vrednih infrastrukturah v EU, ne pa v komercialnih oblakih, kjer obstaja tveganje za rabo podatkov v tržne ali druge nepedagoške namene. Splošna uredba o varstvu podatkov (GDPR) in Akt o umetni inteligenci EU (Akt o umetni inteligenci, 2024) jasno določata, da morajo biti uporabniki seznanjeni s tem, kje se njihovi podatki hranijo, kako se obdelujejo in kdo ima do njih dostop. Tak pristop ne varuje le pravice do zasebnosti, temveč krepi zaupanje v izobraževalni sistem in spodbuja odgovorno digitalno kulturo.

Načelo trajnosti

UI ima tudi svoj ekološki odtis, saj veliki jezikovni modeli za delovanje porabijo ogromne količine energije in čiste vode. Zato mora biti njena uvedba v izobraževanje preiščljena tudi z vidika okoljske in digitalne trajnosti. UNESCO (2023) poudarja, da je trajnost eden ključnih vidikov odgovorne rabe Gen-UI, kar vključuje tako izbiro energetske učinkovitih rešitev kot razmislek o njihovi dejanski pedagoški vrednosti. Zato bi morale šole in univerze spodbujati načelo digitalne umirjenosti (ang. *digital sobriety*), torej zmerno in smiselno uporabo tehnologije, ki sledi resničnim učnim ciljem, ne pa logiki tehnološke novosti ali zgolj učinkovitosti. Tak pristop krepi odgovornost do okolja in hkrati vrača poudarek na bistvo izobraževanja: razvoj razumevanja, kritičnega mišljenja in človeške ustvarjalnosti.

Sklep: med tehnološkim determinizmom in etično avtonomijo

Etični premislek o Gen-UI presega vprašanja pravilne rabe ter posega v bistvo odnosa med človekom in tehnologijo. Na eni strani je tako tehnološki determinizem, povezan s prepričanjem, da razvoja UI ni mogoče zaustaviti, da sledi lastni logiki napredka in da se mu lahko družba le prilagaja. Na drugi strani je etična avtonomija, ki sledi načelni ideji, da mora človek ohraniti nadzor nad smerjo, cilji in pomenom tehnologij, ki jih ustvarja.

Zato je UI v izobraževanju preizkus našega odnosa do lastnega znanja in moči. Če jo razumemo kot orodje, ki naj povečuje človeške sposobnosti, ostajamo v okvirih humanistične tradicije. Če pa postane nadomestek mišljenja in ustvarjalnosti, tvegamo, da se odrečemo prav tistemu, kar nas dela človeške. Gen-UI ni zgolj tehnološki sistem, pač pa tudi proteza, ki razširi naše sposobnosti (Freud, 2021). Toda vsaka proteza spremeni tistega, ki jo uporablja. Tako kot mikroskop razširi naš pogled, a hkrati spremeni način gledanja, tako tudi Gen-UI razširi naš um, a spremeni način mišljenja. Etika v tem kontekstu služi kot orodje samoohranitve človeškega merila, kot varovalka, ki skrbi, da nas moč tehnologije ne razčloveči.

Literatura

- Abrahamson, E. (2025). AI in Higher Education: Bridging the Divide Between Access, Equality, and Opportunity. *International Center for Academic Integrity*. https://academicintegrity.org/aws/ICAI/pt/sd/news_article/587198/_PARENT/layout_interior/false?
- Evropska komisija (2019). *Etične smernice za zaupanja vredno umetno inteligenco*. <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>
- Uredba (EU) 2024/1689 Evropskega parlamenta in Sveta z dne 13. junija 2024 o določitvi harmoniziranih pravil o umetni inteligenci in spremembi uredb (ES) št. 300/2008, (EU) št. 167/2013, (EU) št. 168/2013, (EU) 2018/858, (EU) 2018/1139 in (EU) 2019/2144 ter direktiv 2014/90/EU, (EU) 2016/797 in (EU) 2020/1828 (Akt o umetni inteligenci) (2024). *Uradni list Evropske unije*, L 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Faktabaari (2025). AI guide for teachers. <https://faktabaari.fi/edu/oppaat/teachers-ai/>
- Freud, S. (2021). *Nelagodje v kulturi*. Intelego d.o.o.
- Gonzales, S. (2024). *AI literacy and the new Digital Divide – A Global Call for Action*. UNESCO. <https://www.unesco.org/ethics-ai/en/articles/ai-literacy-and-new-digital-divide-global-call-action>
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I. in Maes, P. (2025). Your Brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay-writing task. *arXiv*. <https://doi.org/10.48550/arXiv.2506.08872>
- Lourdu, V. J., Sawale, P. S., Kaul, P., Pal, S. in Murthy, B. S. R. (2025). Digital divide in AI-powered education: Challenges and solutions for equitable learning. *Journal of Information Systems Engineering and Management*, 10 (21 Suppl), 300–308. <https://doi.org/10.52783/jisem.v10i21s.3327>
- McDermott, B. in Anselmo, L. (2024). Academic Integrity Reflection Tool: Responsible GenAI Use for Students in Assessments. <https://taylorinstitute.ualgary.ca/resources/Academic-Integrity-Reflection-Tool>
- OECD (2023). Emerging governance of generative AI in education. V OECD, *OECD Digital Education Outlook 2023*. https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en/full-report/emerging-governance-of-generative-ai-in-education_3cbd6269.html
- OECD (2024). *OECD AI Principles overview*. <https://oecd.ai/en/ai-principles>
- Nilsson, N. J. (2013). *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge University Press.

- Reich, J. in Dukes, J. (2024). Toward a New Theory of Arrival Technologies. *OSF Preprints*. <https://doi.org/10.35542/osf.io/x6vn7>
- UNESCO (2021). *AI and Education: Guidance for Policy-Makers*. UNESCO. <https://doi.org/10.54675/PCSP7350>
- UNESCO (2023). *Guidance for Generative AI in Education and Research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Xia, Q., Weng, X., Ouyang, F., Zhang, Y. in Li, J. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(40). <https://doi.org/10.1186/s41239-024-00468-z>
- Ženko, E. (2024). Etični vidiki uporabe umetne inteligence v zdravstvu. V T. Štemberger Kolnik in A. Presker Planko (ur.), *Digitalizacija v zdravstvu in umetna inteligenca: inovacije za boljšo prihodnost* (str. 81–93). Fakulteta za zdravstvene vede v Celju.
- Ženko, E. (2025). V znamenju trikotnika: izobraževanje, etika in umetna inteligenca. V A. Flogie in Čotar Konrad, S. (ur.), *Izobraževanje v dobi generativne umetne inteligence: mednarodne smernice in raziskave* (str. 323–344). Založba Univerze na Primorskem. <https://doi.org/10.26493/978-961-293-431-6.16>