



The Holy See

Poslanica svetega očeta Frančiška za 57. svetovni dan miru

1. januar 2024

Umetna inteligenca in mir

Na začetku novega leta, časa milosti, ki ga Gospod podarja vsakemu izmed nas, bi rad nagovoril Božje ljudstvo, narode, voditelje držav in vlad, predstavnike različnih verstev in civilne družbe, vse može in žene našega časa ter jim izrazil svoje voščilo miru.

1. Napredek znanosti in tehnologije kot pot k miru

Sveto pismo potrjuje, da je Bog ljudem dal svojega Duha, da bi imeli »modrost, razumnost in vednost za vsakršno delo« (2 Mz 35,31). Razumnost je izraz dostojanstva, ki nam ga je podaril Stvarnik, ko nas je ustvaril po svoji podobi in podobnosti (prim. 1 Mz 1,26) in nam omogočil, da na njegovo ljubezen odgovorimo s svobodo in spoznanjem. Znanost in tehnologija na poseben način kažeta na to temeljno kakovost odnosne narave človekovega razuma: sta izreden sad njegovih ustvarjalnih sposobnosti.

Drugi vatikanski koncil je v Pastoralni konstituciji *Gaudium et spes* to resnico poudaril in izjavil, da je »človek s svojim delom in duhovno močjo vedno skušal razvijati svoje življenje«. [1] Ko si ljudje »s pomočjo tehnike« prizadevajo, da bi zemlja »postajala dostojno bivališče za celotno človeško družino«, [2] delujejo po Božjem načrtu in sodelujejo z njegovo voljo pri dovršitvi stvarstva in širjenju miru med narodi. Tudi napredek znanosti in tehnike, kolikor prispeva k boljši urejenosti človeške družbe ter k rasti svobode in bratskega občestva, torej vodi k izboljšanju človeka in preobrazbi sveta.

Upravičeno se veselimo in smo hvaležni za izredne dosežke znanosti in tehnologije, ki so pripomogli k odpravi mnogoterega zla, ki je prizadelo človeško življenje in povzročalo veliko trpljenje. Istočasno tehnološko-znanstveni napredek, ki omogoča doslej še neviden nadzor nad resničnostjo, v človekove roke polaga široko paleto možnosti, med katerimi nekatere lahko predstavljajo tveganje za preživetje in nevarnost za skupni dom. [3]

Izjemen napredek novih informacijskih tehnologij, zlasti na digitalnem področju, predstavlja torej navdušujoče priložnosti in resna tveganja s pomembnimi posledicami za dosego pravičnosti in za harmonijo med ljudstvi. Zato si moramo nujno postaviti nekaj vprašanj. Kakšne bodo srednjeročne in dolgoročne posledice novih digitalnih tehnologij? In kakšen vpliv bodo imele na življenje posameznikov in družbe, na mednarodno stabilnost in mir?

2. Prihodnost umetne inteligence med obljubami in tveganji

Napredek informatike in razvoj digitalnih tehnologij v zadnjih desetletjih sta že začela povzročati globoke spremembe v globalnih družbah in v njenih dinamikah. Nova digitalna orodja spreminjajo podobo komunikacij, javne uprave, izobraževanja, potrošnje, medosebnih odnosov in nešteti drugih vidikov vsakdanjega življenja.

Poleg tega lahko tehnologije, ki uporabljajo številne algoritme, iz digitalnih sledi, puščenih na internetu, pridobijo podatke, ki omogočajo nadzor mentalnih in odnosnostnih navad ljudi v komercialne ali politične namene, pogosto brez njihove vednosti, kar omejuje njihovo zavestno udeležanje svobode odločanja. V spletnem prostoru, za katerega je značilna preobremenjenost z informacijami, se dejansko lahko oblikuje tok podatkov glede na izbire, ki jih uporabnik ne zazna vedno.

Ne smemo pozabiti, da znanstvene raziskave in tehnološke inovacije niso vedno »nevtralne«, ločene od stvarnosti, [4] ampak so podvržene kulturnim vplivom. Gre za povsem človeške dejavnosti, zato sprejete usmeritve odražajo odločitve, ki jih pogojujejo osebne, družbene in kulturne vrednote posameznega obdobja. Enako lahko rečemo za rezultate, ki jih dosegajo: kot rezultat specifično človeških pristopov do okolja, v katerem nastajajo, imajo vedno etično razsežnost, ki je tesno povezana z odločitvami tistih, ki načrtujejo poizkuse in usmerjajo proizvodnjo k določenim ciljem.

To velja tudi za oblike umetne inteligence. Do danes v svetu znanosti in tehnologije nimamo enoznačne definicije zanjo. Sam pojem, ki je že stopil v splošno rabo, zajema različne znanosti, teorije in tehnike, ki so usmerjene k temu, da bi stroji pri svojem delovanju reproducirali ali posnemali spoznavne sposobnosti ljudi. Uporaba množine, torej »oblike inteligence« lahko služi temu, da poudarimo predvsem nepremostljivo razliko med temi sistemi, pa naj bodo še tako presenetljivi in vplivni, in človekom: navsezadnje so »fragmentarni«, ker lahko samo posnemajo ali reproducirajo nekatere funkcije človeške inteligence. Uporaba množine poleg tega poudarja tudi,

da je treba te naprave, ki se med seboj zelo razlikujejo, vedno obravnavati kot »družbeno-tehnične sisteme«. Njihov vpliv dejansko poleg temeljne tehnologije ni odvisen samo od načrtovanja, ampak tudi od ciljev in interesov tistega, ki jih ima v lasti in jih razvija, pa tudi od situacij, v katerih se uporabljajo.

Umetno inteligenco moramo torej razumeti kot galaksijo različnih resničnosti in ne moremo vnaprej predvidevati, da bo njen razvoj sam po sebi blagodejno prispeval k prihodnosti človeštva in miru med narodi. Takšen pozitiven izid bo mogoč samo, če bomo dokazali, da smo sposobni delovati odgovorno in spoštovati temeljne človeške vrednote, kot so »vključenost, preglednost, varnost, pravičnost, diskretnost in zanesljivost«. [5]

Poleg tega ni dovolj niti predpostavljati, da bodo tisti, ki načrtujejo algoritme in digitalne tehnologije, ravnali etično in odgovorno. Treba je okrepiti ali, če je to potrebno, oblikovati organizme, pristojne za preučevanje etičnih vprašanj, ki se bodo pojavljala, in za zaščito pravic tistih, ki uporabljajo oblike umetne inteligence ali so pod njenim vplivom. [6]

Izjemen razmah tehnologije mora zato spremljati ustrezna vzgoja k odgovornosti za njen razvoj. Svoboda in mirno sožitje sta ogrožena, kadar ljudje popustijo skušnjavi sebičnosti, osebnega interesa, želji po dobičku in žeji po oblasti. Zato smo dolžni razširiti pogled in tehnično-znanstveno raziskovanje usmeriti k prizadevanju za mir in skupno dobro, k služenju celostnemu razvoju človeka in skupnosti. [7]

Dostojanstvo, ki je lastno vsakemu človeku, in bratstvo, ki nas povezuje kot člane ene same človeške družine, morata biti temelj razvoja novih tehnologij in služiti kot nesporni merili za njihovo ocenjevanje pred njihovo uporabo, da bo digitalni napredek lahko spoštoval pravičnost in prispeval k miru. Tehnološkega razvoja, ki ne vodi k izboljšanju kakovosti življenja vsega človeštva, ampak nasprotno povečuje neenakosti in konflikte, nikoli ne moremo imeti za pravi napredek. [8]

Umetna inteligenca bo postajala vedno bolj pomembna. Njeni izzivi so tako tehnični, kot antropološki, vzgojni, družbeni in politični. Obljublja na primer prihranek napora, učinkovitejšo proizvodnjo, lažji transport in bolj dinamične trge, pa tudi revolucijo v procesih zbiranja, organiziranja in preverjanja podatkov. Zavedati se moramo hitrih sprememb, ki se dogajajo, in jih usmerjati tako, da bomo zavarovali temeljne človekove pravice ob spoštovanju ustanov in zakonov, ki spodbujajo celovit človekov razvoj. Umetna inteligenca mora služiti najboljšim človekovim zmožnostim in našim najvišjim pričakovanjem, ne pa tekmovati z njimi.

3. Tehnologija prihodnosti: stroji, ki se učijo sami

V svojih številnih oblikah umetna inteligenca, ki temelji na tehnikah samodejnega učenja [*machine learning*] in je šele v pionirski fazi, že uvaja izjemne spremembe v tkivo družb, saj globoko vpliva na kulture, družbeno ravnanje in graditev miru.

Razvoj, kot sta strojno učenje in globoko učenje [*deep learning*], odpira vprašanja, ki presegajo področji tehnologije in inženiringa ter segajo na področje razumevanja, ki je tesno povezano s pomenom človeškega življenja, z osnovnimi procesi spoznavanja in sposobnostjo uma, da doseže resnico.

Sposobnost nekaterih naprav, da na primer proizvedejo skladenjsko in pomensko dosledna besedila, ni zagotovilo zanesljivosti. Pravijo, da lahko »halucinirajo«, se pravi, da ustvarjajo trditve, ki se na prvi pogled zdijo sprejemljive, a so v resnici neutemeljene ali izražajo predsodke. To predstavlja resen problem, kadar se umetna inteligenca uporablja v dezinformacijskih kampanjah, ki širijo lažne novice in vodijo k vedno večjemu nezaupanju do medijev. Zaupnost, posedovanje podatkov in intelektualna lastnina so naslednja področja, za katera navedene tehnologije predstavljajo resna tveganja. Tem se pridružujejo negativne posledice, povezane z njihovo neprimerno uporabo, kot so diskriminacija, vmešavanje v volilne procese, krepitev družbe, ki nadzira in kontrolira ljudi, digitalna izključenost, poglobljanje individualizma, ki vedno bolj ločuje od skupnosti. Vsi ti dejavniki predstavljajo tveganje za podžiganje konfliktov in ovirajo mir.

4. Občutek za mejo v tehnokratski paradigmi

Naš svet je preobsežen, raznolik in zapleten, da bi ga lahko v celoti poznali in opredeljevali. Človeški um ne bo mogel nikoli izčrpati njegovega bogastva, tudi s pomočjo najbolj naprednih algoritmov ne. Ti namreč ne zagotavljajo točnih napovedi prihodnosti, ampak samo statistične približke. Vsega ni mogoče napovedati, vsega ni mogoče izračunati; na koncu »resničnost presega idejo« [9] in ne glede na to, kako izredna je lahko naša sposobnost računanja, bo vedno obstajal nek nedosegljiv ostanek, ki se izogne vsakemu poskusu merjenja.

Poleg tega velika količina podatkov, ki jih analizira umetna inteligenca, sama po sebi ni zagotovilo nepristranskosti. Ko algoritmi ekstrapolirajo informacije, so vedno v nevarnosti, da jih popačijo in s tem ponavljajo krivice in predsodke okolij, iz katerih izvirajo. Bolj kot postajajo hitri in zapleteni, težje je razumeti, zakaj so prišli do nekega določenega rezultata.

»Pametni« stroji lahko vse učinkoviteje opravljajo naloge, ki so jim dodeljene, cilj in pomen njihovega delovanja pa bodo še vedno določali ali omogočali ljudje s svojim vesoljem vrednot. Obstaja tveganje, da merila, na katerih temeljijo nekatere izbire, postanejo manj jasna, da se odgovornost za odločanje skrije in da se proizvajalci izognejo obveznosti delovanja v dobro skupnosti. V določenem smislu to podpira tehnokratski sistem, za katerega je gospodarstvo zaveznik tehnologije in ki daje prednost kriteriju učinkovitosti, pri čemer teži k neupoštevanju vsega, kar ni povezano z njegovimi neposrednimi interesi. [10]

To nas mora spodbuditi k razmišljanju o vidiku, ki je v sedanji tehnokratski in učinkovitostni miselnosti zelo pogosto prezrt, a je odločilen za osebni in družbeni razvoj: »občutek za mejo«. Človek, ki je po definiciji umrljiv, a misli, da bo s tehniko prekoračil vsako mejo, je namreč v

nevarnosti, da bo zaradi obsedenosti z željo imeti nadzor nad vsem, izgubil nadzor nad samim seboj; da bo v iskanju absolutne svobode padel v spiralo neke tehnološke diktature. Prepoznanje in sprejetje lastne meje, meje ustvarjenega bitja, je za človeka nujen pogoj za doseg ali bolje sprejetje daru polnosti. Nasprotno pa bi v ideološkem okviru tehnokratske paradigme, ki jo spodbuja prometejska prevzetnost glede samozadostnosti, neenakosti lahko čezmerno narasle, znanje in bogastvo pa bi se kopičila v rokah peščice, z velikim tveganjem za demokratične družbe in za mirno sobivanje. [11]

5. Vroče etične teme

V prihodnosti bi lahko zanesljivost posojilojemalca, primernost posameznika za določeno delo, možnost, da obsojenec ponovi kaznivo dejanje, pravico do političnega azila ali do socialne pomoči določali sistemi umetne inteligence. Pomanjkanje različnih stopenj posredovanja, ki jih ti sistemi uvajajo, je še posebej izpostavljeno različnim oblikam predsodkov in diskriminacije: sistemske napake se zlahka pomnožijo in s tem ne povzročajo samo krivice v posameznih primerih, ampak po učinku verižne reakcije tudi resnične oblike družbene neenakosti.

Poleg tega se včasih zdi, da so oblike umetne inteligence sposobne vplivati na odločitve posameznikov po vnaprej določenih možnostih, ki so povezane s spodbudami in svarili, ali po sistemih za uravnavanje osebnih odločitev, ki temeljijo na organiziranju informacij. Te oblike manipulacije ali družbenega nadzora zahtevajo pozornost in skrben nadzor ter vključujejo jasno zakonsko odgovornost proizvajalcev, uporabnikov in državnih oblasti.

Zanašanje na samodejne procese, ki razvrščajo posameznike, na primer z vsesplošno uporabo nadzora ali s sprejetjem sistemov družbenih bonitetnih ocen, bi lahko imelo globoke učinke tudi na družbeno tkivo in bi vzpostavilo neprimerne lestvice razvrščanja državljanov. Ti umetni procesi razvrščanja bi lahko vodili tudi do konfliktov moči, ki ne bi zadevali samo virtualnih naslovnikov, ampak realne ljudi. Temeljno spoštovanje človekovega dostojanstva zahteva zavrnitev istovetenja enkratnosti osebe zgolj s skupkom podatkov. Algoritmom ne smemo dovoliti, da bi določali, kako razumemo človekove pravice, da bi zanemarjali bistvene vrednote sočutja, usmiljenja in odpuščanja ali da bi odpravili možnost, da se posameznik lahko spremeni in pusti za seboj svojo preteklost.

V tem okviru ne moremo izpustiti presoje vpliva novih tehnologij na področje dela: opravila, ki so bila nekoč izključna domena človeške delovne sile, hitro posrkajo industrijske aplikacije umetne inteligence. Tudi v tem primeru obstaja temeljna nevarnost nesorazmerne koristi za peščico na račun osiromašenja mnogih. V času, ko te oblike tehnologije prodirajo vse globlje na področje dela, bi morale spoštovanje dostojanstva delavcev in pomembnost zaposlitve za ekonomsko blaginjo ljudi, družin in družbe, varnost zaposlitve in pravičnost plač predstavljati visoko prednostno nalogo mednarodne skupnosti.

6. Bomo meče prekovali v lemeže?

Ko v teh dneh gledamo svet okrog nas, ne moremo ubežati resnim etičnim vprašanjem, ki so povezana s panogo oboroževanja. Možnost vodenja vojaških operacij po sistemih za daljinsko vodenje je privedla do manjšega dojemanja opustošenja, ki ga povzročajo, in odgovornosti za njihovo uporabo, kar prispeva k hladnejšemu in še bolj odmaknjenemu odnosu do neizmerne tragedije vojne. Raziskovanje tehnologij, ki se pojavljajo na področju »smrtonosnih avtonomnih oborožitvenih sistemov«, vključno z uporabo umetne inteligence v vojskovanju, so resen razlog za etično zaskrbljenost. Avtomatski oborožitveni sistemi ne bodo mogli biti nikoli moralno odgovorni subjekti: izključno človeška sposobnost moralnega presojanja in etičnega odločanja je več kot zapleten skupek algoritmov in te sposobnosti ne moremo omejiti na programiranje nekega stroja, ki ne glede na to, kako »pameten« je, še vedno ostaja zgolj stroj. Zato je treba nujno zagotoviti primeren, pomenljiv in dosleden človeški nadzor oborožitvenih sistemov.

Prav tako ne smemo prezreti možnosti, da skrajno izpopolnjeno orožje konča v napačnih rokah in omogoči na primer teroristične napade ali posege, ki so namenjeni destabilizaciji zakonitih vladnih ustanov. Svet res nima nobene potrebe, da bi nove tehnologije prispevale k nepravničnemu razvoju trga in trgovine z orožjem ter spodbujale norost vojne. S takšnim ravnanjem ni v nevarnosti le človekova inteligenca, ampak tudi njegovo srce, ki tako postaja vse bolj »umetno«. Najbolj naprednih tehničnih aplikacij ne bi smeli uporabljati za omogočanje nasilnega reševanja konfliktov, ampak za tlakovanje poti miru.

Če pogledamo s pozitivnega vidika, bi umetno inteligenco lahko uporabljali za spodbujanje celostnega človekovega razvoja, kar bi uvedlo pomembne izboljšave v kmetijstvu, izobraževanju in kulturi, izboljšalo življenjsko raven celotnih narodov in ljudstev, omogočilo rast človeškega bratstva in družbenega prijateljstva. Način, kako umetno inteligenco uporabimo za vključevanje najmanjših, to je najslabotnejših in najbolj potrebnih bratov in sester, kaže na stopnjo naše človečnosti.

Človeški pogled in želja po boljši prihodnosti našega sveta vodita do potrebe po interdisciplinarnem dialogu, usmerjenem v etični razvoj algoritmov – k *algor-etiki* –, v kateri bodo vrednote usmerjale poti novih tehnologij. [12] Etična vprašanja bomo morali upoštevati tako na začetku raziskovanja kot tudi v fazi poizkusov, načrtovanja, proizvodnje, distribucije in trženja. Gre za etični pristop k načrtovanju, v katerem imajo bistveno vlogo vzgojne ustanove in odgovorni v procesu odločanja.

7. Izzivi na področju vzgoje

Razvoj tehnologije, ki spoštuje človekovo dostojanstvo in mu služi, ima jasne posledice za vzgojne ustanove in svet kulture. Digitalne tehnologije so s pomnožitvijo možnosti komuniciranja omogočile nove načine srečevanja. Vendarle pa ostaja potreba po nenehnem razmisleku o načinu odnosov,

h katerim nas usmerjajo. Mladi rastejo v kulturnih okoljih, ki so prežeta s tehnologijo, zato se ne moramo ne zamisliti nad metodami poučevanja in izobraževanja.

Vzgoja za uporabo oblik umetne inteligence bi morala biti usmerjena predvsem v spodbujanje kritičnega mišljenja. Nujno je, da uporabniki vseh starosti, predvsem pa mladi, razvijejo sposobnost razločevanja pri uporabi podatkov in vsebin, ki so zbrani na spletu ali jih proizvajajo sistemi umetne inteligence. Šole, univerze in znanstveniki so poklicani, da pomagajo študentom in strokovnjakom usvojiti družbene in etične vidike razvoja ter uporabe tehnologije.

Usposabljanje za uporabo novih komunikacijskih orodij bi moralo upoštevati ne le dezinformacije, lažne novice, ampak tudi zaskrbljujočo razširitev »atavističnih strahov« [...] , ki so se prikriili »in se ob novih tehnologijah okrepili«. [13] Žal se moramo ponovno boriti »s skušnjo po oblikovanju kulture zidov, zidanja zidov, s katerimi bi preprečevali srečanje z drugimi kulturami, z drugim ljudstvom« [14] ter za razvijanje mirnega in bratskega sobivanja.

8. Izzivi za razvoj mednarodnega prava

Globalni doseg umetne inteligence jasno kaže, da imajo poleg odgovornosti suverenih držav za njeno urejanje znotraj sebe tudi mednarodne organizacije odločilno vlogo pri doseganju večstranskih dogovorov ter usklajevanju njihove uporabe in izvajanja. [15] V zvezi s tem spodbujam Skupnost narodov, naj sodeluje pri sprejemanju zavezujoče mednarodne pogodbe, ki bi urejala razvoj in uporabo umetne inteligence v njenih številnih oblikah. Cilj predpisov pa seveda ne bi smel biti samo preprečevanje slabih praks, ampak tudi spodbujanje dobrih, novih in ustvarjalnih pristopov ter omogočanje osebnih in kolektivnih pobud. [16]

Končno, v iskanju normativnih modelov, ki razvijalcem digitalnih tehnologij lahko nudijo etične smernice, je nujno določiti človeške vrednote, ki morajo biti v temelju prizadevanja družbe za oblikovanje, sprejetje in uporabo potrebnih zakonodajnih okvirov. Delo za urejanje etičnih smernic za proizvodnjo oblik umetne inteligence ne more zanemariti upoštevanja globljih vprašanj, ki zadevajo pomen človeškega bivanja, varovanje temeljnih človekovih pravic, prizadevanje za pravičnost in mir. Ta proces etičnega in pravnega razločevanja se lahko izkaže kot dragocena priložnost za skupen razmislek o vlogi, ki bi jo tehnologija morala imeti v življenju posameznika in skupnosti, in o tem, kako njena uporaba lahko prispeva k ustvarjanju bolj pravičnega in človeškega sveta. Zato bi morali v razpravah o urejanju umetne inteligence upoštevati glas vseh strani, ki jih to zadeva, vključno z ubogimi, obrobni in drugimi, ki v globalnih procesih odločanja pogosto ostajajo neslišani.

* * * * *

Upam, da nas bo to razmišljanje spodbudilo k skrbi, da bo napredek v razvoju oblik umetne inteligence služil človeškemu bratstvu in miru. Ne gre samo za odgovornost nekaterih, ampak za

odgovornost celotne človeške družine. Mir je namreč sad odnosov, ki priznavajo in sprejmejo bližnjega v njegovem neodtujljivem dostojanstvu, ter sodelovanja in zavzetosti pri iskanju celostnega razvoja vseh ljudi in vseh ljudstev.

Na začetku novega leta molim, da hiter razvoj oblik umetne inteligence ne bi povečal v svetu že tako prevelikih neenakosti in krivic, ampak bi prispeval h končanju vojn in spopadov ter olajšanju številnih oblik trpljenja, ki pestijo človeško družino. Naj verni kristjani, verniki različnih verstev ter moške in žene dobre volje usklajeno sodelujejo, da bodo izkoristili priložnosti in se spoprijeli z izzivi, ki jih prinaša digitalna revolucija, ter prihodnjim rodovom izročili bolj solidaren, pravičen in miroljuben svet.

Vatikan, 8. decembra 2023

FRANČIŠEK

[1] Št. 33.

[2] Prav tam, št. 57.

[3] Prim. Okrožnica *Hvaljen, moj Gospod* (24. maja 2015), št. 104.

[4] Prim. *Prav tam*, št. 114.

[5] *Avdiencia za udeležence srečanja "Minerva Dialogues"* (27. marca 2023).

[6] Prim. *Prav tam*.

[7] Prim. *Pismo izvršnemu predsedniku „World Economic Forum“ v Davos-Klosters* (12. januarja 2018).

[8] Prim. Okrožnica *Hvaljen, moj Gospod*, št. 194; Nagovor udeležencem na seminarju *Skupno dobro v digitalni dobi* (27. septembra 2019).

[9] Apostolska spodbuda *Veselite evangelija* (24. novembra 2013), št. 233.

[10] Prim. Okrožnica *Hvaljen, moj Gospod*, št. 54.

[11] Prim. Nagovor udeležencem plenarnega zasedanja Papeške akademije za življenje (28. februarja 2020).

[12] Prim. *Prav tam*.

[13] Okrožnica *Vsi smo bratje* (3. oktobra 2020), št. 27.

[14] Prim. *Prav tam*.

[15] Prim. *Prav tam*, št. 170–175.

[16] Prim. Okrožnica *Hvaljen, moj Gospod*, št. 177.