



The Holy See

PORUKA SVETOGA OCA
FRANJE
ZA
57. SVJETSKI DAN MIRA

1. SIJEČNJA 2024.

Umjetna inteligencija i mir

Na početku nove godine, milosnog vremena koje Gospodin daruje svakome od nas, želim se obratiti narodu Božjem, narodima, čelnicima država i vlada, predstavnicima različitih vjera i civilnog društva, te svim muškarcima i ženama našeg vremena kako bih im uputio najbolje želje za mir.

1. Napredak znanosti i tehnologije kao put do mira

Sveto pismo svjedoči da je Bog ljudima dao svoga Duha kako bi ih resila »umješnost, sposobnost i razumijevanje u svim poslovima« (Izl 35, 31). Inteligencija je izraz dostojanstva kojim nas je obdario Stvoritelj, koji nas je stvorio na svoju sliku i priliku (usp. Post 1, 26) i osposobio nas da na njegovu ljubav odgovorimo slobodom i spoznajom. Znanost i tehnologija na osobit način pokazuju tu fundamentalno relacijsku prirodu ljudske inteligencije: oni su izvanredan plod njezina kreativnog potencijala.

U Pastoralnoj konstituciji *Gaudium et spes* Drugi vatikanski koncil potvrdio je tu istinu izjavivši: »Svojim radom i snagom svojega duha čovjek je uvijek nastojao što šire razviti svoj život« [1]. Kad ljudi nastoje »pomoću tehničkih umijeća« Zemlju učiniti »dostojnim prebivalištem cjelokupne ljudske obitelji« [2], oni djeluju po Božjem naumu i surađuju s njegovom voljom da dovrše djelo stvaranja i šire mir među narodima. I napredak znanosti i tehnologije, u mjeri u kojoj pridonosi boljem uređenju ljudskog društva, rastu i razvoju slobode i bratskog zajedništva, vodi, dakle, čovjekovu boljitku i preobrazbi svijeta.

S pravom se radujemo i zahvalni smo za izvanredna dostignuća znanosti i tehnologije zahvaljujući kojima se doskočilo bezbrojnim zlima koja su mučila ljudski život i uzrokovala veliku patnju. Istodobno, znanstveni i tehnološki napredak, koji omogućuje dosad nezapamćenu kontrolu nad stvarnošću, stavlja ljudima u ruke mnoštvo mogućnosti od kojih neke mogu predstavljati prijetnju čovjekovom opstanku i opasnost za zajednički dom [3].

Značajni napreci u novim informacijskim tehnologijama, posebno na digitalnom polju, donose, dakle, zadivljujuće mogućnosti i ozbiljne opasnosti, s ozbiljnim implikacijama na težnju za postizanjem pravde i sklada među narodima. Zato je prijeko potrebno postaviti neka hitna pitanja. Koje će, srednjoročno i dugoročno gledano, biti posljedice novih digitalnih tehnologija? I kakav će utjecaj one imati na život pojedinaca i društva, na međunarodnu stabilnost i mir?

2. Budućnost umjetne inteligencije između obećanja i rizika

Napreci postignuti u informacijskoj tehnologiji i razvoj digitalnih tehnologija proteklih desetljeća već su doveli do dubokih promjena u globalnom društvu i njegovoj dinamici. Novi digitalni alati mijenjaju lice komunikacije, javne uprave, obrazovanja, potrošnje, osobnih interakcija i bezbrojnih drugih vidova svakodnevnog života.

Nadalje, pomoću tehnologija koje koriste pregršt algoritama moguće je iz digitalnih tragova ostavljenih na internetu izvući podatke koji omogućuju kontrolu razmišljanja i navika ljudi u odnosima, često bez njihovog znanja, u komercijalne ili političke svrhe, čime se ograničava njihovu svjesnu slobodu izbora. U internetskom prostoru prenakrcanom informacijama oni mogu oblikovati tok podataka prema kriterijima odabira koje korisnik ne zamjećuje uvijek.

Moramo zapamtiti da znanstveno istraživanje i tehnološke inovacije nisu odvojene od stvarnosti i »neutralne« [4], nego su podložne kulturnim utjecajima. Kako je riječ o potpuno ljudskim aktivnostima, smjerovi kojim idu odražavaju izbore uvjetovane osobnim, društvenim i kulturnim vrijednostima svakog doba. Isto vrijedi i za rezultate koje postižu: upravo zato što su plod specifičnog ljudskog pristupa svijetu koji nas okružuje, uvijek imaju etičku dimenziju, usko povezanu s odlukama onih koji provode eksperimente i usmjeravaju proizvodnju prema određenim ciljevima.

To vrijedi i za oblike umjetne inteligencije. U svijetu znanosti i tehnologije zasad ne postoji jedinstvena definicija za nju. Sam izraz, koji je već ušao u svakodnevnu upotrebu, obuhvaća niz znanjâ, teorijâ i tehnikâ kojima je cilj postići to da strojevi, u svom radu, reproduciraju ili oponašaju kognitivne sposobnosti ljudi. Govoriti u množini o "oblicima inteligencije" može nadasve pomoći da se naglasi nepremostivi jaz koji postoji između tih sustava, ma koliko nevjerojatni i moćni bili, i ljudske osobe: oni su, u konačnici, "fragmentarni", u smislu da mogu samo oponašati ili reproducirati neke funkcije ljudske inteligencije. Upotrebom množine također se naglašava da te, međusobno jako različite, uređaje treba uvijek promatrati kao "sociotehničke sustave". Naime,

njihov učinak, neovisno o temeljnoj tehnologiji, ne ovisi samo o načinu na koji su izrađeni, nego i o ciljevima i interesima onih koji ih posjeduju i onih koji ih razvijaju, te o situacijama u kojima se koriste.

Umjetnu inteligenciju stoga treba shvaćati kao zaseban svijet sazdan od različitih stvarnosti pa ne možemo *a priori* pretpostaviti da će njezin razvoj dati pozitivan doprinos budućnosti čovječanstva i miru među narodima. Takav pozitivan ishod bit će moguć samo ako se pokažemo sposobnima ponašati se odgovorno i poštivati temeljne ljudske vrijednote kao što su »inkluzivnost, transparentnost, sigurnost, pravda, povjerljivost i pouzdanost« [5].

Nije dovoljno, isto tako, pretpostaviti da će se oni koji razvijaju algoritme i digitalne tehnologije ponašati etično i odgovorno. Potrebno je ojačati ili, ako je potrebno, uspostaviti tijela zadužena za razmatranje novih etičkih pitanja i zaštitu pravâ onih koji koriste neki od oblika umjetne inteligencije ili su izloženi njihovom utjecaju [6].

Silno veliko širenje tehnologije mora stoga biti praćeno odgovarajućom izgradnjom za preuzimanje odgovornosti za njezin razvoj. Sloboda i miran suživot su ugroženi kad ljudi podlegnu iskušenju egoizma, osobnog interesa, žudnje za profitom i želje za moći. Imamo stoga dužnost proširiti svoj pogled i usmjeriti tehnološka i znanstvena istraživanja prema težnji za postizanjem mira i općeg dobra, u službi cjelovitog razvoja čovjeka i zajednice [7].

Neotuđivo dostojanstvo svakog ljudskog bića i bratstvo koje nas povezuje kao članove jedne ljudske obitelji moraju biti u temelju razvoja novih tehnologija i služiti kao neosporan kriterij za njihovu procjenu prije nego ih se stavi u upotrebu, kako bi se digitalni napredak ostvarivao u duhu pravde i pridonosio miru. Tehnološki razvoji koji ne vode do poboljšanja kvalitete života cijelog čovječanstva, nego naprotiv pogoršavaju nejednakosti i sukobe, nikako se ne mogu smatrati pravim napretkom [8].

Umjetna inteligencija postajat će sve važnija. Izazovi koje pred nas stavlja su tehnički, ali i antropološki, didaktički, društveni i politički. Obećava, primjerice, uštedu naporâ, efikasniju proizvodnju, učinkovitiji transport i dinamičnija tržišta kao i revoluciju u prikupljanju, organizaciji i provjeri podataka. Moramo biti svjesni brzih promjena koje se trenutno događaju i njima upravljati na način da se štiti temeljna ljudska prava i poštuje institucije i zakone koji promiču cjelovit ljudski razvoj. Umjetna inteligencija mora biti u službi najboljih ljudskih potencijala i naših najviših ciljeva, a ne u suparništvu s njima.

3. Tehnologija budućnosti: strojevi koji uče sami

U svojim mnogostrukim oblicima, umjetna inteligencija, utemeljena na tehnikama strojnog učenja (*machine learning*), premda je još u pionirskoj fazi, već uvodi značajne promjene u strukturu društava i ima dubok utjecaj na kulture, društvena ponašanja i izgradnju mira.

Razvoji kao što su strojno učenje (*machine learning*) ili duboko učenje (*deep learning*) pokreću pitanja koja nadilaze područja tehnologije i inženjerstva i tiču se razumijevanja koje je usko povezano sa smislom ljudskog života, temeljnim kognitivnim procesima i sposobnošću uma da se vine do istine.

Tako, primjerice, sposobnost nekih uređaja da iznjedre sintaktički i semantički suvisle tekstove nije jamstvo pouzdanosti. Za njih se kaže da mogu "halucinirati", odnosno generirati tvrdnje koje se na prvi pogled čine uvjerljivima, ali zapravo su neutemeljene ili odaju predrasude. To predstavlja ozbiljan problem kad se umjetna inteligencija koristi u kampanjama dezinformiranja koje šire lažne vijesti i vode do rastućeg nepovjerenja u sredstva komunikacije. Povjerljivost, vlasništvo nad podacima i intelektualno vlasništvo druga su područja u kojima te tehnologije za sobom povlače ozbiljne opasnosti, kojima valja pribrojati druge negativne posljedice njihove zlouporabe, kao što su diskriminacija, uplitanje u izborne procese, nastanak modela društva koje nadzire i kontrolira ljude, digitalna isključenost i jačanje individualizma sve više otrgnutog od kolektiviteta. Svi ti čimbenici prijete da raspiruju sukobe i da budu prepreka miru.

4. Osjećaj ograničenosti u tehnokratskoj paradigmi

Naš je svijet pregolem, previše raznolik i presložen da bi ga se moglo u potpunosti upoznati i klasificirati. Ljudski um nikada neće moći iscrpiti sve njegovo bogatstvo, pa ni uz pomoć najnaprednijih algoritama. Ovi potonji, naime, ne nude nikakva pouzdana predviđanja za budućnost, već samo statističke aproksimacije. Ne može se sve predvidjeti, ne može se sve izračunati. U konačnici, »stvarnost [...] je važnija od ideje« [9] i bez obzira na to koliko naši računalni kapaciteti bili čudesni, uvijek će postojati nedostupan ostatak koji izmiče svim pokušajima mjerenja.

K tome, velika količina podataka koje analizira umjetna inteligencija sama po sebi nije jamstvo nepristranosti. Kad algoritmi ekstrapoliraju informacije, uvijek se izlažu opasnosti da ih iskrive, replicirajući nepravde i predrasude okruženja iz kojeg potječu. Što su brži i složeniji, to je teže razumjeti zašto su dali određeni rezultat.

"Inteligentni" strojevi mogu obavljati dodijeljene im zadaće sa sve većom učinkovitošću, ali svrhu i značenje njihovih operacija i dalje će određivati ili omogućavati ljudi, svaki sa svojim svijetom vrijednosti. Tu se javlja opasnost da kriteriji na kojima se temelje određene odluke postanu manje jasni, da se odgovornost za odluke zamagljuje i da proizvođači izbjegavaju svoju obvezu djelovanja za dobrobit zajednice. To, u stanovitom smislu, potiče tehnokratski sustav koji „stvara savez“ između ekonomije i tehnologije i daje povlašteno mjesto kriteriju učinkovitosti, težeći tome da zanemari sve što nije povezano s njegovim neposrednim interesima [10].

To nas mora potaknuti na razmišljanje o vidu koji je, u današnjem tehnokratskom mentalitetu usmjerenom na učinkovitost, tako često zanemaren, a ipak je ključan za osobni i društveni razvoj,

a to je “osjećaj ograničenosti”.

Kad čovjek, koji je po definiciji smrtan, misli da tehnologijom može nadići svaku ograničenost, u opasnosti je da, opsjednut željom da sve kontrolira, izgubi kontrolu nad samim sobom; da, u potrazi za apsolutnom slobodom, upadne u spiralu tehnološke diktature. Prepoznati i prihvatiti vlastitu ograničenost kao stvorenja neizostavan je uvjet da čovjek postigne puninu ili, bolje rečeno, prihvati puninu kao dar. Međutim, u ideološkom kontekstu tehnokratske paradigme, nadahnute prometejskom pretpostavkom samodostatnosti, nejednakosti bi mogle rasti preko mjere, a znanje i bogatstvo mogli bi se gomilati u rukama nekolicine, što za sobom povlači ozbiljne opasnosti za demokratska društva i miran suživot [11].

5. *Goruća pitanja za etiku*

U budućnosti bi sustavi umjetne inteligencije mogli određivati pouzdanost zajmotražitelja, prikladnost osobe za neki posao, vjerojatnost da će neki osuđenik ponoviti zlodjelo ili pravo na dobivanje političkog azila ili socijalne pomoći. Nepostojanje različitih razina posredovanja koje ti sustavi uvode posebno je izloženo određenim oblicima predrasude i diskriminacije: pogreške u sustavu mogu se lako umnožiti i dovesti ne samo do nepravde u pojedinačnim slučajevima, već i, domino efektom, do stvarnih oblika društvene nejednakosti.

Kadikad se, k tome, čini da oblici umjetne inteligencije mogu utjecati na odluke pojedinaca kroz unaprijed određene opcije povezane s poticajima i razuvjeravanjima (odvrćanjima) odnosno putem sustavâ koji reguliraju osobne izbore temeljene na organizaciji podataka. Ti oblici manipulacije ili društvene kontrole zahtijevaju pažljivu pozornost i nadzor i za sobom povlače jasnu zakonsku odgovornost od strane proizvođačâ, korisnikâ i državnih vlasti.

Oslanjanje na automatizirane procese koji kategoriziraju pojedince, na primjer sveprisutnom upotrebom sustava nadzora ili uvođenjem sustava društvenog rejtinga, također bi moglo imati nesagledive posljedice na društvo, uspostavljajući neprikladna rangiranja građanâ. A ti umjetni procesi klasifikacije mogli bi također dovesti do sukobâ moći, jer se ne tiču samo virtualnih primatelja, već i stvarnih ljudi. Temeljno poštivanje ljudskog dostojanstva zahtijeva odbacivanje poistovjećivanja jedinstvenosti osobe sa skupom podataka. Ne smije se dopustiti da algoritmi diktiraju način na koji shvaćamo ljudska prava, da se ostavi po strani temeljne vrjednote suosjećanja, milosrđa i oprostjenja ili ukloni mogućnost da se pojedinac može promijeniti i ostaviti prošlost iza sebe.

U tom pogledu naprosto moramo razmišljati o utjecaju novih tehnologija na svijet rada: poslove koji su nekoć bili isključiva domena radne snage brzo preuzimaju industrijske primjene umjetne inteligencije. I u ovom slučaju postoji velika opasnost od stjecanja nerazmjerne koristi nekolicine na račun osiromašivanja mnogih pojedinaca. Poštivanje dostojanstva radnikâ i važnost zapošljavanja za materijalno blagostanje pojedinaca, obitelji i društava, sigurnost poslova i

pravedne plaće trebali bi biti visoki prioritet za međunarodnu zajednicu, dok ti oblici tehnologije sve dublje prodiru u svijet rada.

6. Hoćemo li mačeve prekovati u plugove?

U današnje vrijeme, gledajući svijet koji nas okružuje, nemoguće je izbjeći ozbiljna etička pitanja vezana uz sektor naoružanja. Mogućnost vođenja vojnih operacija korištenjem daljinski upravljanih sustava dovela je do smanjene percepcije razaranja koja za sobom ostavljaju i odgovornosti za njihovo korištenje, pridonoseći još hladnijem i distanciranijem odnosu prema golemoj tragediji rata. Istraživanje tehnologija koje se javljaju na području takozvanih “smrtonosnih autonomnih oružanih sustava”, uključujući korištenje umjetne inteligencije u ratu, ozbiljan je razlog etičke zabrinutosti. Autonomni oružani sustavi nikada neće moći biti moralno odgovorni subjekti: isključivo ljudska sposobnost moralnog prosuđivanja i etičkog odlučivanja više je od složenog skupa algoritama a ta se sposobnost ne može svesti na programiranje stroja koji, ma koliko “inteligentan” bio, uvijek ostaje stroj. Zato jamčenje odgovarajućeg, smislenog i dosljednog ljudskog nadzora nad oružanim sustavima predstavlja imperativ.

Ne možemo zanemariti ni mogućnost da sofisticirano oružje dospije u pogrešne ruke, omogućavajući, na primjer, terorističke napade ili intervencije sa ciljem destabilizacije legitimnih sustava vlasti. Ukratko, svijetu doista nisu potrebne nove tehnologije koje pridonose nepravednom razvoju tržišta i trgovine oružjem, promičući tako bezumlje rata. Na taj način, ne samo ljudskoj inteligenciji, već i samom čovjekovu srcu, prijete opasnost da postane sve više “umjetno”. Najnaprednije tehničke aplikacije ne bi se trebale koristiti za olakšavanje nasilnog rješavanja sukobâ, već za utiranje putova miru.

Gledano s pozitivnijeg stanovišta, kad bi se umjetnu inteligenciju koristilo za promicanje cjelovitog ljudskog razvoja, to bi moglo dovesti do važnih inovacija u poljoprivredi, obrazovanju i kulturi, poboljšanju životnog standarda cijelih nacija i naroda te rastu ljudskog bratstva i društvenog prijateljstva. Na kraju krajeva, način na koji je koristimo da uključimo najmanje, odnosno našu najslabiju i braću i sestre koji su u najvećoj potrebi, mjera je koja pokazuje našu ljudskost.

Humana perspektiva i želja za boljom budućnošću našeg svijeta dovode do potrebe za interdisciplinarnim dijalogom usmjerenim na etički pristup razvoju algoritama – *algoretiku* – gdje će vrjednote biti te koje će usmjeravati putove razvoja novih tehnologija [12]. Etička pitanja treba uzimati u obzir od početka istraživanja, kao i u fazama testiranja, razvoja, proizvodnje, distribucije i komercijalizacije. To je pristup etike planiranja u kojem obrazovne ustanove i odgovorni za proces donošenja odluka imaju ključnu ulogu.

7. Izazovi za obrazovanje

Razvoj tehnologije koja poštuje i služi ljudskom dostojanstvu ima jasne implikacije za obrazovne

ustanove i svijet kulture. Digitalne su tehnologije umnožavanjem mogućnosti komunikacije omogućile nam susretati se na nove načine. No, potrebno je neprestano razmišljati o vrsti odnosa prema kojima nas usmjeravaju. Mladi odrastaju u kulturnim okruženjima prožetima tehnologijom, što neizbježno za sobom povlači pitanja vezana uz metode nastave i obrazovanja.

Podučavanje korištenju raznih oblika umjetne inteligencije prvenstveno treba imati za cilj promicanje kritičkog mišljenja. Prijeko je potrebno da korisnici svih dobnih skupina, a posebno mladi, razviju sposobnost kritičkog korištenja podataka i sadržaja kojima se pristupa putem interneta ili koje kreiraju sustavi umjetne inteligencije. Škole, sveučilišta i znanstvene zajednice pozvane su pomoći studentima i stručnjacima da ovladaju društvenim i etičkim vidovima razvoja i korištenja tehnologije.

U odgoju za korištenje novih komunikacijskih sredstava trebalo bi uzeti u obzir ne samo dezinformaciju, *fake news*, već i uznemirujuće ponovno oživljavanje »iskonskih strahova [koji] su se uspjeli sakriti i ojačati u okrilju novih tehnologija« [13]. Ponovno se, nažalost, suočavamo s »napašću da se gradi kulturu zidova, da se podižu zidovi, zidovi u srcu, zidovi na zemlji, kako bi se spriječio ovaj susret s drugim kulturama, s drugim ljudima« [14] i razvoj mirnog i bratskog suživota.

8. Izazovi za razvoj međunarodnoga prava

Globalni doseg umjetne inteligencije jasno pokazuje da, osim odgovornosti suverenih država da reguliraju njezinu uporabu unutar svojih granica, međunarodne organizacije mogu igrati ključnu ulogu u sklapanju multilateralnih sporazuma te u koordinaciji njihove primjene i provedbe [15]. U vezi s tim, pozivam međunarodnu zajednicu na zajednički rad u cilju usvajanja obvezujućeg međunarodnog ugovora kojim će se regulirati razvoj i korištenje umjetne inteligencije u njezinim različitim oblicima. Ta bi regulacije, naravno, trebala imati za cilj ne samo sprječavanje štetnih praksi, već i poticanje dobrih praksi, poticanjem novih i kreativnih pristupa i olakšavanjem pojedinačnih i grupnih inicijativa [16].

U konačnici, u potrazi za normativnim propisima koji mogu pružiti etičke smjernice onima koji rade na razvijanju digitalnih tehnologija, bitno je prepoznati i utvrditi ljudske vrijednosti na kojima bi se trebali temeljiti napori društava oko formuliranja, usvajanja i primijene potrebnih pravnih okvira. U radu na izradi etičkih smjernica za proizvodnju oblika umjetne inteligencije ne može se zanemariti dublja pitanja o smislu ljudskog postojanja, zaštiti temeljnih ljudskih prava i težnji za postizanjem pravde i mira. Taj proces etičkog i pravnog razlučivanja može se pokazati dragocjenom prigodom za zajedničko promišljanje o ulozi koju bi tehnologija trebala imati u našim životima kao pojedinaca i zajednice kao i o tome kako njezino korištenje može pridonijeti stvaranju pravednijeg i humanijeg svijeta. Zato se u raspravama o regulaciji umjetne inteligencije treba uzeti u obzir glasove svih zainteresiranih strana, uključujući siromašne, marginalizirane i druge koji su u globalnim procesima donošenja odluka često bez glasa.

Nadam se da će nas ova razmišljanja potaknuti da se pobrinemo da napredak u razvoju oblika umjetne inteligencije u konačnici služi cilju ljudskog bratstva i mira. To nije odgovornost nekolicine, već čitave ljudske obitelji. Mir je, naime, plod odnosâ koji drugog prepoznaju i prihvaćaju u njegovu neotuđivom dostojanstvu kao i suradnje i predanog zalaganja u potrazi za cjelovitim razvojem svih ljudi i svih naroda.

Moja molitva na početku nove godine jest ova: da brzi razvoj oblika umjetne inteligencije ne poveća nejednakosti i nepravde kojih je već ionako premnogo u svijetu, nego da pomogne u okončanju ratova i sukoba i ublažavanju mnogih oblika patnji koji muče ljudsku obitelj. Neka kršćani, vjernici različitih religija i muškarci i žene dobre volje, skladno surađuju kako bi iskoristili prilike i uhvatili se u koštac sa izazovima koje pred nas stavlja digitalna revolucija i predali budućim naraštajima svijet u kojem će biti više solidarnosti, pravde i mira.

Iz Vatikana, 8. prosinca 2023.

FRANJO

[1] Br. 33.

[2] *Isto*, br. 57.

[3] Usp. enc. *Laudato si'* (24. svibnja 2015.), 104.

[4] Usp. *isto*, 114.

[5] *Obraćanje sudionicima skupa „Minerva Dijalozi” (Udienza ai partecipanti all’Incontro “Minerva Dialogues”)* (27. ožujka 2023.).

[6] Usp. *isto*.

[7] *Poruka predsjedavajućem “Svjetskog ekonomskog foruma” u Davos-Klostersu (Messaggio al Presidente Esecutivo del “World Economic Forum” a Davos-Klosters)* (12. siječnja 2018.).

[8] Usp. enc. *Laudato si'*, 194; *Obraćanje sudionicima seminara “Opće dobro u digitalnom dobu” (Discorso ai partecipanti al Seminario “Il bene comune nell’era digitale”)* (27. rujna 2019.).

[9] Apost. pob. *Evangelii gaudium* (24. studenog 2013.), 233.

[10] Usp. enciklika *Laudato si'*, 54.

[11] Usp. *Obraćanje sudionicima plenarne skupštine Papinske akademije za život (Discorso ai partecipanti alla Plenaria della Pontificia Accademia per la Vita)* (28. veljače 2020.).

[12] Usp. *isto*.

[13] Enc. *Fratelli tutti* (3. listopada 2020.), 27.

[14] Usp. *isto*.

[15] Usp. *isto*, 170-175.

[16] Usp. enc. *Laudato si'*, 177.