

Značajke etike umjetne inteligencije

Hergouth-Vicencinović, Andre Roman

Undergraduate thesis / Završni rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka, Faculty of Tourism and Hospitality Management / Sveučilište u Rijeci, Fakultet za menadžment u turizmu i ugostiteljstvu**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:191:364157>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-02-06**



Repository / Repozitorij:

[Repository of Faculty of Tourism and Hospitality Management - Repository of students works of the Faculty of Tourism and Hospitality Management](#)



SVEUČILIŠTE U RIJECI
FAKULTET ZA MENADŽMENT U TURIZMU I UGOSTITELJSTVU
OPATIJA

ANDRÉ ROMAN HERGOUTH-VICENCINOVIĆ

ZNAČAJKE ETIKE UMJETNE INTELIGENCIJE
CHARACTERISTICS OF ETHICS IN ARTIFICIAL
INTELLIGENCE

ZAVRŠNI RAD

RIJEKA, 2023.

SVEUČILIŠTE U RIJECI
FAKULTET ZA MENADŽMENT U TURIZMU I UGOSTITELJSTVU
OPATIJA

ANDRÉ ROMAN HERGOUTH-VICENCINOVIĆ

ZNAČAJKE ETIKE UMJETNE INTELIGENCIJE
CHARACTERISTICS OF ETHICS IN ARTIFICIAL
INTELLIGENCE

ZAVRŠNI RAD

Kolegij: Etika i društvena odgovornost

Mentor: Prof. dr. sc. Marija Ivaniš

Student: André Roman Hergouth-Vicencinović

Matični broj: ps25125mo19

Smjer: MOR

RIJEKA, svibanj 2023.



SVEUČILIŠTE U RIJECI UNIVERSITY OF RIJEKA
FAKULTET ZA MENADŽMENT U TURIZMU I UGOSTITELJSTVU
FACULTY OF TOURISM AND HOSPITALITY MANAGEMENT
OPATIJA, HRVATSKA CROATIA

IZJAVA O AUTORSTVU RADA I O JAVNOJ OBJAVI RADA

Ime i prezime studenta: André Roman Hergouth Vicencinović

Matični broj: ps25125mo19

Izjavljujem da sam završni rad pod naslovom

Značajke etike umjetne inteligencije

(Naslov rada)

izradila/o samostalno te sam suglasna/suglasan o javnoj objavi rada u elektroničkom obliku.

Svi dijelovi rada, nalazi ili ideje koje su u radu citirane ili se temelje na drugim izvorima, bilo da su u pitanju knjige, znanstveni ili stručni članci, Internet stranice, zakoni i sl. u radu su jasno označeni kao takvi te adekvatno navedeni u popisu literature.

U Opatiji, _____ 2023.

Potpis studenta

SADRŽAJ:

SAŽETAK

1. UVOD	1
1.1. Svrha i cilj istraživanja	1
1.2. Predmet, problematika i objekti istraživanja	1
1.3. Hipoteza i pomoćne hipoteze	2
1.4. Ocjene dosadašnjih istraživanja	2
1.5. Struktura rada	3
2. UMJETNA INTELIGENCIJA – TEORIJSKA I POJMOVNA RAZMATRANJA	4
2.1. Etičke smjernice poduzeća u razvijanju umjetne inteligencije	4
2.1.1. Karakteristike etike umjetne inteligencije	5
2.1.2. Etika i moral poduzeća koja razvijaju umjetnu inteligenciju	5
2.2. Preporuke smjernica etike umjetne inteligencije	6
2.2.1. Smjernice UNESCO-a.....	6
2.2.2. Smjernice OECD-a.....	7
3. VAŽNOST I ULOGA UMJETNE INTELIGENCIJE U SUVREMENOM SUSTAVU POSLOVANJA	9
3.1. Problematika donošenja pravila o implementaciji umjetne inteligencije	9
3.1.1. Donošenje pravila i normi o etici umjetne inteligencije.....	9
3.1.2. Stavovi korisnika o umjetnoj inteligenciji.....	10
3.2. Problematika etike umjetne inteligencije u zdravstvenom sustavu	10
3.2.1. Etičke barijere zdravstvenog sustava.....	11
3.2.2. Obrada podataka pacijenata.....	11
3.3. Primjena umjetne inteligencije u poslovnoj praksi na primjeru Alibaba Grupe	11
4. MEĐUZAVISNOST TRANSPARENTNOSTI, ETIKE I ODGOVORNOSTI	13
4.1. Značajke vjerodostojnosti umjetne inteligencije	13
4.1.1. Temeljni zahtjevi umjetne inteligencije	13
4.1.2. Problematika vjerodostojnih sistema umjetne inteligencije na primjeru Michaela Millera.....	13
4.2. Glavne debate etike umjetne inteligencije	15
4.2.1. Privatnost i nadzor.....	15
4.2.2. Manipulacija ponašanja	16
4.2.3. Autonomija sistema.....	17
4.2.4. Prozirnost.....	17
4.2.5. Singularnost.....	18

4.2.6.	Automatizacija i zapošljavanje	19
4.2.7.	Pri stranost odlučivanja	19
5.	ZAKLJUČAK	21
LITERATURA		

SAŽETAK

Razvoj i inovacije u svijetu su sve brže i kompleksnije naravni. Kompleksnost je u većinskom slučaju sinonim za efikasnost i poboljšanje predmeta u pitanju. Porastom produktivnosti i općenito ekonomskog stanja, dobiva se na moći i utjecaju te se implementacijom etičkih načela suzbija negativni utjecaj koji bi ona mogla imati na nama već poznate načine kojih smo svjedoci kroz povijest ljudskog napretka. Kada su u pitanju kompleksne inovacije, umjetna inteligencija je definitivno jedna od najkompleksnijih. Njezina dubina kompleksnosti je nama još daleko od poznatog, nadalje čovjek se tek upoznaje sa njezinim pozitivnim i negativnim stranama. Trenutno je većina polja djelatnosti u čistoj teoriji što se tiče implementacije i stvaranja funkcionalnog sustava umjetne inteligencije.

Etika kao znanost se bavi proučavanjem ljudskog ponašanja te se temelji na moralnim vrijednostima. Etika je kroz povijest enormno dobila na značaju, sukobi i ishodi koji su se događali i još se događaju, proizvod su čvrste specifično orijentirane etike određenog skupa ljudi te se međusobni sporazumi različitih skupina teško ostvaruju zbog navedenog. Sami značaj etike može se vidjeti iz nastalih pojmova poput poslovne etike, odgovornog menadžmenta, itd. Ovi pojmovi obuhvaćaju cjelokupno ponašanje u poslovanju s naglašenim moralnim vrijednostima i obzirom prema prirodi. U konstantno razvijajućim granama poslovanja i zemljama u kojima su poduzeća stacionirana, implementacija etičkih smjernica donosi dugoročno pozitivni tijek djelovanja na sveukupno poslovanje interne i eksterne okoline. Dok u nekim slabo razvijenim zemljama to nije opcija zbog slabog poticaja i infrastrukture, utjecaj globalizacije svakako stvara pritisak na nj te potencira promjenu.

Ovaj rad prolazi kroz etičke teme i nužne tehnološke stavke vezane uz umjetnu inteligenciju i njen način rada. Teorijski i praktično ćemo prikazati primjer djelovanja umjetne inteligenciju u današnjoj ekonomskoj infrastrukturi. Proći ćemo pozitivne i negativne utjecaje, izazove unutar razvoja i ostale dodatne teorijske teme.

Ključne riječi: etika, umjetna inteligencija, suvremeno poslovanje.

1. UVOD

Uvod kao jedan od najbitnijih dijelova završnog rada, upućuje čitatelja na koji je način rad koncipiran:

1. Svrha i cilj istraživanja
2. Problematika i predmet istraživanja
3. Hipoteza i pomoćne hipoteze
4. Znanstvene metode istraživanja
5. Struktura rada

1.1. Svrha i cilj istraživanja

SVRHA ISTRAŽIVANJA je pojasniti smjernice razvoja etike umjetne inteligencije te njihove prijetnje i prilike u ekonomskom, ekološkom i društvenom okruženju.

CILJ ISTRAŽIVANJA ovog završnog rada je stvoriti sliku pravca međusobne prilagodbe čovjeka i tehnologije (umjetne inteligencije).

Pojedinačni ciljevi:

1. Stvoriti svijest o prijetnjama i prilikama koje se pojavljuju pojavom umjetne inteligencije.
2. Prikazati ustanovljene granice i preporuke za razvoj daljnjih ograničenja i načina pristupa budućem razvoju umjetne inteligencije.
3. Prikazati trenutnu implementaciju umjetne inteligencije u određenim poslovnim sustavima.

1.2. Predmet, problematika i objekti istraživanja

Pojava i razvoj umjetne inteligencije budi interes u ekonomskom svijetu. Prednosti koje potencira umjetna inteligencija su izrazite, ali jednako tako stvara i značajne prijetnje najviše društvu i kulturi. Određivanje smjernica i ograničenja razvoja etike umjetne inteligencije je od kritičnog značaja za kvalitetu i korist koju umjetna inteligencija može pružiti čovjekovu blagostanju.

PROBLEM ISTRAŽIVANJA je nejasni pravac razvoja umjetne inteligencije i određivanje ograničenja razvoja i pristupa.

PREDMET ISTRAŽIVANJA ovog završnog rada je pravilna interpretacija značajki etike umjetne inteligencije i pristup prema njihovom razvoju i implementaciji.

OBJEKTI ISTRAŽIVANJA su: etika, umjetna inteligencija, suvremeno poslovanje..

Ključne riječi: etika, umjetna inteligencija, suvremeno poslovanje.

1.3. Hipoteza i pomoćne hipoteze

Postavlja se sljedeća **Glavna Hipoteza**: Prepoznaje se problematika i izazov vezan uz uspješno implementiranje i planiranje etičnosti unutar sustava umjetne inteligencije u sustave određenih poslovnih grana.

Glavna hipoteza je postavljena uz pomoć tri pomoćne hipoteze:

H1: Upoznava se etika i preporučene smjernice postavljene od različitih tijela u korist boljeg razumijevanja problematike vezane uz umjetnu inteligenciju.

H2: Proučavaju se problematika i uspjeh vezani uz implementaciju etične umjetne inteligencije unutar zdravstvenog i poslovnog sektora potrebno za daljnje razmatranje i planiranje tijeka i kvalitete razvoja umjetne inteligencije.

H3: Analiza vjerodostojnosti sustava i pitanja moralnih vrijednosti su nužni za etičnu umjetnu inteligenciju.

1.4. Ocjene dosadašnjih istraživanja

Brojni autori su svoje radove namjenili istraživanju etike umjetne inteligencije. Postojeća istraživanja nam prikazuju razne perspektive i problematiku vezanu navedenu temu.

Dušić, Tea (2021.) – Kroz svoj rad, Etika umjetne inteligencije¹, govori o tehnologiji sadašnjosti i njenoj sve bržoj implementaciji u raznim sektorima. Diskutira o negativnim i pozitivnim stranama implementacije umjetne inteligencije i etičkim načelima vezanih uz samu tehnologiju i ponašanje stručnjaka. Završno pitanje koje postavlja je: Koristimo li tehnologiju na ispravan način, odnosno je li ona zaista dobra?

Puzek, Željka (2018.) – Svojim radom, Etika umjetne inteligencije², dotiče se tema razvoja umjetne inteligencije i njenih mogućih ishoda. Rizici, najvažniji trenutci i etičke implikacije su predstavljene kroz ovaj rad te mogući načini prevencije kako bi se minimizirale mogućnosti negativnih posljedica. Krajnje se utvrđuje uloga i odgovornost etike unutar ovog složenog procesa.

Jaklić, Valentina (2019.) – Opisuje kroz svoj rad, Utjecaj umjetne inteligencije u poslovanju³, umjetnu inteligenciju i njen utjecaj na poslovanje, povijesni razvoj, etiku robotike i smjerove u istraživanju. Obrađuje teme vezane uz inovativnost i snagu promjene koju ima umjetna inteligencija, navodi kako se brzina razvijanja i širenja umjetne inteligencije prkosi sa dostupnom literaturom te ne mogu pratiti uspješnost proizvoda umjetne inteligencije. Naglašava kako način življenja neće biti kao do sada, barem kod tehnologije.

¹ T. Dušić, "Etika umjetne inteligencije", Diplomski rad, Sveučilište u Rijeci, Filozofski fakultet, Rijeka, 2021. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:186:018413>

² Ž. Puzek, "Etika umjetne inteligencije", Diplomski rad, Sveučilište u Zagrebu, Filozofski fakultet, Zagreb, 2018. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:131:795477>

³ V. Jaklić, "Utjecaj umjetne inteligencije u poslovanju", Diplomski rad, Sveučilište u Zagrebu, Ekonomski fakultet, Zagreb, 2019. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:148:363157>

1.5. Struktura rada

Završni rad sadrži pet tematski povezanih dijelova.

U **UVODU** završnog rada određuju se svrha i ciljevi istraživanja, nadalje predmet, problematika i objekti istraživanja. Postavlja se glavna hipoteza uz pomoćne hipoteze, određuju se znanstvene metode koje su korištene te se opisuje struktura završnog rada.

Drugi dio završnog rada, **UMJETNA INTELIGENCIJA – TEORIJSKA I POJMOVNA RAZMATRANJA**, sastoji se od sljedećih cjelina: *1.) Etičke smjernice poduzeća u razvijanju umjetne inteligencije, 2.) Preporuke smjernica etike umjetne inteligencije*, od kojih svaka sadrži svoje podcjeline. U prvoj cjelini se izlaže stav i moral određenih poduzeća u vezi razvijanja i dijeljenja svojih usluga sa ostalim strankama te se iznose karakteristike etike umjetne inteligencije. Druga cjelina prikazuje preporuke i smjernice razvijanja etike umjetne inteligencije po UNESCO-u, i OECD-u. Većinom teorijski pristup određivanju smjernica i ograničenja razvitka umjetne inteligencije.

Treći dio završnog rada, **VAŽNOST I ULOGA UMJETNE INTELIGENCIJE U SUVREMENOM SUSTAVU POSLOVANJA**, sastoji se od: *1.) Problematika donošenja pravila o implementaciji umjetne inteligencije, 2.) Problematika etike umjetne inteligencije u zdravstvenom sustavu, 3.) Primjena umjetne inteligencije u poslovnoj praksi*. Prisutne su podcjeline unutar prve dvije cjeline. U prvoj cjelini je glavna tema donošenje pravila o implementaciji umjetne inteligencije između države i poduzeća. Druga cjelina prikazuje trenutnu problematiku unutar zdravstvenog sustava upravljanog putem umjetne inteligencije. Treća cjelina prikazuje primjere poslovne primjene umjetne inteligencije u praksi.

Četvrti dio, **MEĐUZAVISNOST TRANSPARENTNOSTI, ETIKE I ODGOVORNOSTI**, sastoji se o: *1.) Značajke vjerodostojnosti umjetne inteligencije, 2.) Glavne debate etike umjetne inteligencije*. Svaka cjelina ima svoje podcjeline. Prva cjelina iznosi načela i temeljne potrebe umjetne inteligencije za prihvatljivu i moralno upotrebljivu korist ekonomiji i društvu. Druga cjelina iznosi glavne teorijske debate o etici umjetne inteligencije vezane uz prijetnje društvu.

Peti i posljednji dio završnog rada je **ZAKLJUČAK**. On predstavlja sažetak i konačni stav subjekata o etici umjetne inteligencije.

2. UMJETNA INTELIGENCIJA – TEORIJSKA I POJMOVNA RAZMATRANJA

Umjetna inteligencija je novodošla tehnologija. Cjelokupni spektar njene primjene je još u velikoj količini čista teorija. Određivanjem željene crte kretanja razvitka umjetne inteligencije, postavlja se pravac prilagodbe čovjeka i tehnologije.

Stoga se u ovom poglavlju razmatraju sljedeće tematske cjeline: 1.) Etičke smjernice poduzeća u razvijanju umjetne inteligencije, 2.) Preporuke smjernica etike umjetne inteligencije.

2.1. Etičke smjernice poduzeća u razvijanju umjetne inteligencije

Kako bi se sigurnost rukovanja i korištenja visoko napredne tehnologije kao što je umjetna inteligencija stabilizirala i držala pod kontrolom, poduzeća kreiraju svoju politiku i etiku pružanja usluga vezanih uz isto.

U daljnjem tekstu ove cjeline proučit će se etičke smjernice poduzeća C3.ai smještenog u SAD-u⁴.

Poduzeće C3.ai jasno iznosi kako „zakoni i regulative generalno nisu dovoljni kako bi se osiguralo etičko korištenje umjetne inteligencije. Dužnost je pojedinca i organizacija koje koriste umjetnu inteligenciju, kao i poduzeća koji razvijaju i nude usluge povezane uz umjetnu inteligenciju da prakticiraju etiku umjetne inteligencije. Korisnici i Dobavljači umjetne inteligencije moraju poduzeti proaktivne korake kako bi osigurali etičko korištenje umjetne inteligencije. Ova obveza nadilazi izdavanje izvjava, moraju postojati posebne politike koje se aktivno provode“⁵.

Daljnje izjavljuju kako „poduzeće C3.ai prakticira etičku umjetnu inteligenciju. C3.ai vjeruje da su integritet i etičko ponašanje ključni za uspješno poslovanje. Bavimo se stvaranjem poslovnih aplikacija umjetne inteligencije za tvrtke, vlade te javne i privatne organizacije koja im pomaže da dramatično poboljšaju svoje poslovanje i svoje proizvode za dobrobit čovječanstva“⁶.

Održavanje kontrole moguće je putem postavljanja etičkih elemenata, navode „Naša predanost etičkoj umjetnoj inteligenciji proteže se na način na koji dopuštamo našim klijentima da koriste našu umjetnu inteligenciju“⁷.

Njihovi elementi etičkih pravila o umjetnoj inteligenciji uključuju; „1.) Radit ćemo sa snažnim demokratskim vladama, 2.) Nećemo dopustiti totalitarnim državama da koriste našu umjetnu inteligenciju kako bi unaprijedile svoje režime ili kontrolirale svoje građane, 3.) Nećemo dopustiti da se naša umjetna inteligencija koristi za obmanjivanje ljudi ili za dodjeljivanje vrijednosti putem društvenog bodovanja i drugih sustava koje pokreće umjetna inteligencija, 4.) Radit ćemo u energetske industriji na poboljšanju učinkovite proizvodnje i isporuke energije svih vrsta, uključujući obnovljive, 5.) Ako saznamo da se naša umjetna inteligencija koristi na načine koji krše našu etiku, isključit ćemo prekidač“⁸.

⁴ Dostupno na: www.3.ai/glossary/artificial-intelligence/ethical-ai/ (20.4.2023.)

⁵ ibidem

⁶ ibidem

⁷ ibidem

⁸ ibidem

Navedeno poduzeće izjasnilo je svoje etičke vrijednosti i kulturu poslovanja u cilju dobiti čovjeka i društva.

2.1.1. Karakteristike etike umjetne inteligencije

Primjer karakteristika etike umjetne inteligencije obraditi će se na primjeru poduzeća Brighterion. Prema članku Georgea Lawtona „AI ethics (AI code of ethics)“, Sudhir Jha je kazao kako etički kodeks umjetne inteligencije njegova poduzeća prati sljedeća načela kako bi bio razvijen; „Etički sistem umjetne inteligencije mora biti inkluzivan, objašnjiv, pozitivne koristi i mora vršiti obradu podataka na odgovoran način“⁹.

Daljnje objašnjenje navodi kako; „1.) Inkluzivni sistem umjetne inteligencije je nepristran i funkcionira kroz cijeli spektar društva. To zahtjeva cjelovito znanje svakog pojedinog izvora podataka korištenog za obuku modela umjetne inteligencije kako bi se osiguralo da u skupu podataka nema inherentne pristranosti. Također zahtjeva pažljivu reviziju uvježbanog modela kako bi se filtrirali svi problematični atributi naučeni u procesu. Modeli bi se trebali pažljivo motriti kako ne bi došlo do korupcije u budućnosti. 2.) Objašnjivi sistem umjetne inteligencije podupire vladanje zahtijevano od poduzeća kako bi se osiguralo etičko korištenje umjetne inteligencije. Teško je biti siguran u djela umjetne inteligencije koja ne može biti objašnjena. Stjecanje povjerenja može dovesti do kompromisa u kojem se mali kompromis performanse modela učini kako bi se odabrao algoritam koji je objašnjiv. 3.) Sustav umjetne inteligencije koji ima pozitivnu svrhu ima za cilj, na primjer: smanjiti prijevare, eliminirati otpad, nagraditi ljude, usporiti klimatske promjene, izliječiti bolesti itd. Bilo koja se tehnologija može koristiti za nanošenje štete, ali je imperativ da smislimo načine da se zaštiti umjetna inteligencija od iskorištavanja u loše svrhe. Ovo će biti težak izazov, ali s obzirom na širok opseg i razmjere umjetne inteligencije, rizik od nerješavanja ovog izazova i zlouporabe ove tehnologije daleko je veći nego ikad prije. 4.) Sustav umjetne inteligencije koji odgovorno koristi podatke poštuje prava na privatnost podataka. Podaci su ključni za sustav umjetne inteligencije, a više podataka često rezultira boljim modelima. Međutim ključno je da u utrci za prikupljanjem sve više i više podataka ne bude žrtvovano pravo ljudi na privatnost i transparentnost. Odgovorno prikupljanje, upravljanje i korištenje podataka ključno je za stvaranje sustava umjetne inteligencije kojem se može vjerovati. U idealnom svijetom podaci bi se trebali prikupljati samo kada su potrebni, a ne kontinuirano, a granularnost podatak trebala bi biti što je moguće uža. Na primjer, ako aplikacija treba samo geo-lokacijske podatke na razini poštanskog broja za pružanje vremenske prognoze, ne bi trebala prikupljati točnu lokaciju korisnika. A sustav bi trebao rutinski brisati podatke koji više nisu potrebni“¹⁰.

2.1.2. Etika i moral poduzeća koja razvijaju umjetnu inteligenciju

Ostvarenje tehnološkog i moralnog napretka kroz umjetnu inteligenciju dolazi s raznim prijetnjama. Veliku konkurentsku prednost koju nudi umjetna inteligencija, lagano se iskoristi te usmjeri u nemoralne načine implementacije. Kao i sa svakom novom tehnologijom, iskustva je nedovoljno za realno predviđanje i očekivanje ishoda, nadalje traži se izuzetni oprez pri rukovanju s tehnologijom velikog potencijala.

⁹ Dostupno na: <https://www.techtarget.com/whatis/definition/AI-code-of-ethics> (27.04.2023.)

¹⁰ ibidem

Poduzeća su svjesna negativnih utjecaja na razvoj umjetne inteligencije od strane raznih moćnika našeg ekonomskog svijeta. Pojedina poduzeća zasigurno naginju prema novčanoj kompenzaciji za manjak morala, dok se većina poduzeća u demokratskim državama drži moralnih smjernica i općenito etičkog postupanja oko razvoja i implementacije potencijala umjetne inteligencije.

Sive zone korištenja veliki su izazov za etičko ponašanje umjetne inteligencije. Problematika tehnologije vezana uz obradu i sveobuhvatno tumačenje podataka te razne druge debate je jedan od najvećih izazova za razvitak etičke umjetne inteligencije. Neke od glavnih debate vezanih uz etičko ponašanje umjetne inteligencije koje ćemo proučiti u četvrtoj cjelini, obradio je Vincent C. Müller (Ethics of Artificial Intelligence and Robotics)¹¹.

Odgovornost u određenim granicama o ponašanju sustava umjetne inteligencije preuzima poduzeće koje ga razvija. Prateći legislative i smjernice dobivene od nadležnih tijela, poduzeća imaju odgovornost razviti vjerodostojan sustav umjetne inteligencije. Smjernice se razvijaju od strane različitih visokopozicioniranih administracija i organizacija u različitim granama tržišta, jer svako tržište koje je u mogućnosti koristiti sustav umjetne inteligencije, ima svoj niz izazova vezanih uz implementaciju i etičko ponašanje tog sustava.

2.2. Preporuke smjernica etike umjetne inteligencije

Kako bi se ostvario sklad ekonomije i blagostanja društva između država i poduzeća, potrebna je diskusija i iznošenje stavova o planovima razvitka i primjene umjetne inteligencije te etike koju je neizbježno uspostaviti.

U sljedećem tekstu, sagledati će se preporučene smjernice za razvoj etike umjetne inteligencije od strane UNESCO-a i OECD-a.

2.2.1. Smjernice UNESCO-a

UNESCO kao organizacija koja se zalaže za međusobno razumijevanje i poštovanje, koristi edukaciju, kulturu, komunikaciju, znanost i informacije kako bi doprinijela razvoju globalnog društva.

Pojavom umjetne inteligencije, stvara se novi predmet koji treba biti usmjeren prema pozitivnom djelovanju. Prijetnje i prednosti koje se ukazuju kroz razvitak umjetne inteligencije su značajne te ih je potrebno ograničiti i usmjeriti.

UNESCO objavljuje svoje preporuke za etiku umjetne inteligencije „Recommendations on the Ethics of Artificial Intelligence“¹² 2022. godine kojima se konstatira željeni razvojni tijek umjetne inteligencije.

Uvodno se razrađuje problematika povezana uz interakciju čovjeka i umjetne inteligencije te posljedice interakcije. Naglasak je na važnosti etike umjetne inteligencije, uvažavajući potencijalni

¹¹ Müller, Vincent C., "Ethics of Artificial Intelligence and Robotics", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N, 2021.

¹² UNESCO, Recommendations on the Ethics of Artificial Intelligence, SHS/BIO/REC-AIETHICS/2021, 2021.

utjecaj na blagostanje individualaca i društva. Preambula uvažava potencijalne benefite umjetne inteligencije u promoviranju održivog razvoja i čovjekova blagostanja, ali također uzima u obzir moguće rizike, poput prijetnje prema privatnosti, sigurnosti i ljudskoj autonomiji. Naglašava potrebu za orijentacijom umjetne inteligencije prema čovjeku, promovirajući etička načela poput transparentnosti, odgovornosti i inkluzivnosti, kao i važnost interdisciplinarnе kolaboracije i uključenosti dionika. Preambula poziva na kolektivni međunarodni napor za razvijanje smjernica i politike koje osiguravaju odgovorni razvoj i korištenje umjetne inteligencije.

Opseg primjene navodi kako se „preporuke za etičku problematiku u domeni umjetne inteligencije odnose do razine mandata UNESCO-a“¹³.

Nadalje „Preporuke nemaju ambiciju pružiti jednu jedinu definiciju umjetne inteligencije, jer bi se takva definicija mijenjala kroz vrijeme, u skladu razvitka tehnologije“¹⁴. Pruža se posebna pažnja širim etičkim implikacijama sistema umjetne inteligencije u odnosu sa centralnim domenama UNESCO-a.

2.2.2. Smjernice OECD-a

Organizacija za ekonomsku suradnju i razvoj stvorila je pet kategorija smjernica preporuka za donošenje političkih odluka vezanih uz etiku umjetne inteligencije: „1.) Investiranje u istraživanje i razvoj umjetne inteligencije, 2.) Poticanje digitalnog ekosustava za umjetnu inteligenciju, 3.) Pružanje poticajnog političkog okruženja za umjetnu inteligenciju, 4.) Izgradnja ljudskih kapaciteta i priprema za tranziciju tržišta rada, 5.) Međunarodna suradnja za pouzdanu umjetnu inteligenciju“¹⁵.

Vezano uz prvu kategoriju smjernica, donesene su sljedeće smjernice; „1.) Vlade bi trebale razmotriti dugoročna javna ulaganja i poticati privatna ulaganja u istraživanje i razvoj, uključujući interdisciplinarnе napore, kako bi potaknule inovacije u pouzdanoj umjetnoj inteligenciji koja je usredotočena na izazovna tehnička pitanja i društvene pravne i etičke implikacije i političke probleme povezane s umjetnom inteligencijom. 2.) Vlade bi također trebale razmotriti javna ulaganja i poticati privatna ulaganja u otvorene skupove podataka koji su reprezentativni i poštuju privatnost i zaštitu podataka kako bi podržali okruženje za istraživanje i razvoj umjetne inteligencije bez neprikladnih pristranosti te poboljšali interoperabilnost i korištenje standarda“¹⁶.

Vezano uz drugu kategoriju smjernica, donesena je sljedeća smjernica; „Vlade bi trebale poticati razvoj i pristup digitalnom ekosustavu za pouzdanu umjetnu inteligenciju. Takav ekosustav posebno uključuje digitalne tehnologije i infrastrukturu te mehanizme za dijeljenje znanja o umjetnoj inteligenciji, prema potrebi. S tim u vezi, vlade bi trebale razmotriti promicanje mehanizama, poput povjerenja podataka, za podršku sigurnoj, pravednoj, zakonitoj i etičkoj razmjeni podataka“¹⁷.

Treća kategorija smjernica, donosi sljedeće smjernice; „1.) Vlade bi trebale promicati političko okruženje koje podržava agilni prijelaz iz faze istraživanja i razvoja u fazu implementacije i rada za pouzdane sustave umjetne inteligencije. U tu svrhu trebali bi razmotriti korištenje eksperimentiranja kako bi osigurali kontrolirano okruženje u kojemu se sustavi umjetne inteligencije mogu testirati i, prema potrebi, povećati. 2.) Vlade bi trebale pregledati i prilagoditi, prema potrebi,

¹³ ibidem

¹⁴ ibidem

¹⁵ Dostupno na: www.oecd.ai/en/ai-principles (22.04.2023.)

¹⁶ Dostupno na: <https://oecd.ai/en/dashboards/ai-principles/P10> (22.04.2023.)

¹⁷ Dostupno na: <https://oecd.ai/en/dashboards/ai-principles/P11> (22.04.2023.)

svoju politiku i regulatorne okvire i mehanizme procjene koji se primjenjuju na sustave umjetne inteligencije kako bi potaknule inovacije i natjecanje za pouzdanu umjetnu inteligenciju¹⁸.

Četvrta kategorija smjernica, donosi sljedeće smjernice; „1.) Vlade bi trebale blisko surađivati s dionicima kako bi se pripremile za transformaciju svijeta rada i društva. Oni bi trebali osnažiti ljude za učinkovitu upotrebu i interakciju s AI sustavima u cijelom nizu aplikacija, uključujući opremanje potrebnim vještinama. 2.) Vlade bi trebale poduzeti korake, uključujući kroz socijalni dijalog, kako bi osigurale pravednu tranziciju za radnike kako se AI primjenjuje, kao što su programi osposobljavanja tijekom radnog vijeka, potpora onima koji su pogođeni emigracijom i pristup novim prilikama na tržištu rada. 3.) Vlade bi također trebale blisko surađivati s dionicima kako bi promicale odgovornu upotrebu umjetne inteligencije na poslu, kako bi poboljšale sigurnost radnika i kvalitetu radnih mjesta, kako bi potaknule poduzetništvo i produktivnost te nastojale osigurati široku i pravednu raspodjelu koristi od umjetne inteligencije¹⁹.

Peta i zadnja kategorija smjernica, donosi sljedeće smjernice; „1.) Vlade, uključujući zemlje u razvoju i dionike, trebale bi aktivno surađivati kako bi unaprijedile ova načela i napredovale u odgovornom upravljanju pouzdanom umjetnom inteligencijom. 2.) Vlade bi trebale surađivati u OECD-u i drugim globalnim i regionalnim forumima kako bi potaknule razmjenu znanja o umjetnoj inteligenciji, prema potrebi. Oni bi trebali poticati međunarodne, međusektorske i otvorene inicijative s više dionika za prikupljanje dugoročne stručnosti o umjetnoj inteligenciji. 3.) Vlade bi trebale promicati razvoj konsenzusom vođenih globalnih tehničkih standarda s više dionika za interoperabilnu i pouzdanu umjetnu inteligenciju. 4.) Vlade bi također trebale poticati razvoj i vlastitu upotrebu međunarodno usporedivih metrika za mjerenje istraživanja, razvoja i implementacije umjetne inteligencije te prikupiti bazu dokaza za procjenu napretka u provedbi ovih načela²⁰.

¹⁸ Dostupno na: <https://oecd.ai/en/dashboards/ai-principles/P12> (22.04.2023.)

¹⁹ Dostupno na: <https://oecd.ai/en/dashboards/ai-principles/P13> (22.04.2023.)

²⁰ Dostupno na: <https://oecd.ai/en/dashboards/ai-principles/P14> (22.04.2023.)

3. VAŽNOST I ULOGA UMJETNE INTELIGENCIJE U SUVREMENOM SUSTAVU POSLOVANJA

Ekonomija je pokretač razvoja gospodarstva. Za napredak ekonomije potrebni su nam razni resursi poput; znanja, energije, tehnologije, prirodnih resursa, ljudi itd.

Inovacija i brzina kao najveća konkurentna prednost u tržišnoj utakmici, glavni je motiv razvoja novih tehnologija. Umjetna inteligencija kao jedan od trenutno najperspektivnijih inovacija za stjecanje konkurentnosti, donosi izuzetni potencijal za multipliciranjem profita, unošenje novih perspektiva u poslovnim sustavima, dosezanje zacrtanih ciljeva održivosti itd. Isto tako je prisutan potencijal za negativnim efektima i posljedicama ukoliko se razvoj, implementiranje i planiranje ne provede uz dodatnu pažnju i potrebno testiranje s naglaskom na etiku umjetne inteligencije.

U ovoj cjelini, obrađuju se teme vezane uz implementaciju i trenutne probleme funkcionalnosti etike umjetne inteligencije u poslovnim sustavima.

Cjeline koje se obrađuju su: 1.) Problematika donošenja pravila o implementaciji umjetne inteligencije, 2.) Problematika etike umjetne inteligencije u zdravstvenom sustavu, 3.) Primjena umjetne inteligencije u poslovnoj praksi na primjeru Alibaba Grupe.

3.1. Problematika donošenja pravila o implementaciji umjetne inteligencije

Država kao zakonodavna vlast, donosi regulative, zakone i ostale pravne odredbe vezane za teritorij svoje vladavine. Nužnost donošenja pravnih akata pritišće politiku za intervenciju te predstavlja rizik nedovoljnog poznavanje teme i subjekta povezanih uz te akte. Stručne preporuke i zdrave prakse se često ne slažu sa političkim interesima stranaka i država, nadalje razumijevanje i pronalazak bolje opcije povezane uz korist i implementaciju subjekta nije efikasan zbog nedovoljnog znanja i razumijevanja subjekta te općenito novog polja tehnologije.

Zadužene su sposobne institucije i organizacije za stvaranje smjernica i preporuka o etici sustava. Izvršnu i daljnje oblikovnu vlast ima država te se preporučene smjernice oblikuju po potrebi i većem interesu.

3.1.1. Donošenje pravila i normi o etici umjetne inteligencije

Poduzetništvo kao korisnik sustava umjetne inteligencije prati određene smjernice o implementaciji. Standarde i preporuke etičkih načela sustava umjetne inteligencije donose razne organizacije i institucije sposobne za takvu prosudbu.

Naime poduzeća stvaraju sami svoje principe etičkog ponašanja sustava. Stvaranjem svojih principa, kupci i partneri mogu procijeniti kako će voditi svoj odnos s određenim poduzećem. Poduzeća su dužna navesti i dati do znanja svojim korisnicima kako će se njihovi podaci koristiti. Ukoliko se to ne učini ili prekrši i pritaji, određene službe će se zauzeti protiv takvog postupanja (npr. GDPR). Ključni principi GDPR-a su: „1.) pravedna i zakonita obrada; 2.) ograničenje namjene; 3.) minimizacija podataka i zadržavanje podataka“²¹.

²¹ Europska komisija, *Javna uprava i zaštita podataka, Koji su glavni aspekti Opće uredbe o zaštiti podataka (GDPR) kojih bi javna uprava trebala biti svjesna?:* <https://commission.europa.eu/law/law-topic/data-protection/reform/rules->

Bijela kuća u svom „Bill of Rights“²² navodi pet principa vodilja za dizajn, korištenje i implementaciju sustava za američko društvo: „1.) Sigurni i efektivni sustavi, 2.) Obrana od algoritamske diskriminacije, 3.) Privatnost podataka, 4.) Obavijest i obrazloženje, 5.) Ljudske alternative, razmatranja i povratak“²³.

Microsoft je stvorio svojih šest principa za odgovornu umjetnu inteligenciju: „1.) Odgovornost, 2.) Uključenost, 3.) Pouzdanost i sigurnost, 4.) Pravednost, 5.) Transparentnost, 6.) Privatnost i zaštita“. Navedeni principi su vođeni dvjema perspektivama; Etičnost i objašnjivost²⁴. Ovakvi principi su nužni kako se umjetna inteligencija širi u sve više proizvoda i usluga.

3.1.2. Stavovi korisnika o umjetnoj inteligenciji

Prema istraživanjima „Redpoint Global“²⁵, vodeće software-opskrbljujuće tvrtke, sagledavaju se stavovi korisnika o korištenju umjetne inteligencije i strojnog učenja, kako bi se poboljšalo njihovo iskustvo.

„Istraživanje je pokazalo da 77 posto potrošača vjeruje da pozitivna korisnička iskustva i dalje trebaju element ljudskog dodira, a 58 posto potrošača kaže da žele da tvrtke budu jasne kada se koristi umjetna inteligencija. Interakcija između ljudi i umjetne inteligencije također mora biti besprijekorna“.

Nadalje, „Potrošači rangiraju „chatbotove“ kao najidealniju upotrebu umjetne inteligencije za poboljšanje korisničkog iskustva, ali 70 posto preferira ljudsku interakciju u odnosu na njih“.

„Generacijski, udobnost s AI znatno varira. Generacija Z i Milenijalci imaju više povjerenja u način na koji robne marke koriste AI nego Generacija X i Boomer.“

„Postoji nekoliko pozitivnih učinaka koje potrošači ističu kao važne u korištenju umjetne inteligencije, na primjer, 54 posto potrošača smatra da umjetna inteligencija može pomoći robnim markama da osobne informacije budu sigurnije i sigurnije“

3.2. Problematika etike umjetne inteligencije u zdravstvenom sustavu

Zdravstveni sustav je kompleksan po pitanju funkcionalnosti etičke obrade podataka. Uvođenjem nadogradnje sustava putem umjetne inteligencije, pojavljuju se raznoliki problemi i izazovi povezani sa etičkim ponašanjem i etičkom funkcionalnosti sustava.

[business-and-organisations/public-administrations-and-data-protection/what-are-main-aspects-general-data-protection-regulation-gdpr-public-administration-should-be-aware_en#references](https://www.whitehouse.gov/ostp/ai-bill-of-rights/) (23.04.2023.)

²² The White House, *Blueprint for AI Bill of Rights*: www.whitehouse.gov/ostp/ai-bill-of-rights/ (23.04.2023.)

²³ ibidem

²⁴ Microsoft, *Responsible and trusted AI*: <https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/innovate/best-practices/trusted-ai> (23.04.2023.)

²⁵ Businesswire, *News*: www.businesswire.com/news/home/20230131005091/en/73-of-Consumers-Believe-AI-can-have-a-Positive-Impact-on-their-Customer-Experience (23.04.2023.)

3.2.1. Etičke barijere zdravstvenog sustava

Napredak zdravstva može se ostvariti kroz razvoj sustava umjetne inteligencije. Potencijal koji nudi implementacija umjetne inteligencije više je nego privlačan, nadalje kompleksan. Ljudi kao glavni subjekt etičke koristi takvog sustava, moraju biti u središtu pažnje kada se priča o sigurnosti i negativnim posljedicama koje bi mogle proizaći kroz nepažljivu implementaciju istog.

Pod sigurnost spadaju razni faktori koji su dio čovjekova blagostanja. Prava na slobodu, privatnost i puno drugih, dolaze u raznim oblicima ovisno o kakvom području je riječ.

Neke od glavnih etičkih barijera proći će se detaljnije u četvrtom poglavlju ovog završnog rada.

3.2.2. Obrada podataka pacijenata

Podaci su u velikoj količini povjerljive naravi te iz tog razloga predstavljaju izazov implementacije etike umjetne inteligencije u sustav i načina obrade tih podataka. Kako se ne bi prekršila prava pacijenata i kako bi sve ostalo u kontroliranim uvjetima, uzima se veliki oprez pri dizajniranju i implementaciji modela umjetne inteligencije.

Zdravstveni sustavi trpe sve veći pritisak zbog rasta potražnje pacijenata, manjka resursa i slično. Sustavi zdravstva se sve više digitaliziraju i optimiziraju te su podaci u sve većem broju prisutni. Ti podaci bi mogli, ukoliko se pravilno koriste i tumače, pomoći u dijagnozama, praćenju tijeka bolesti i u još puno raznih procesa. Iz tih razloga, pri stvaranju regulativa o umjetnoj inteligenciji, zakonodavna tijela i ostali dionici moraju uzeti u obzir potencijal koji se nalazi u umjetnoj inteligenciji kako bi se ostvario veliki inovativni skok u zdravstvu.

„Korištenje nove tehnologije izaziva zabrinutost oko mogućnosti da će postati novi izvor netočnosti i povrede podataka. U visokorizičnom području zdravstva, pogreške mogu imati teške posljedice za pacijenta koji je žrtva tih pogrešaka. To je ključno zapamtiti jer pacijenti dolaze u kontakt s kliničarima kada su u najranjivijim vremenima“²⁶.

Opasnosti obrade podataka kroz umjetnu inteligenciju zazivaju na izuzetan oprez. Povjerljivost i privatnost podataka je područje skrivene prijetnje i izloženosti. Društvo se u širem spektru pokazuje nesvjesno negativni koja može proizaći kroz digitalne otiske koje ostavlja.

3.3. Primjena umjetne inteligencije u poslovnoj praksi na primjeru Alibaba Grupe

Primjena i prednosti umjetne inteligencije obrađuje se kroz primjer Alibaba Grupe.

²⁶ Helen Smith, Clinical AI: opacity, accountability, responsibility and liability. (2020) 36:535–45. doi: 10.1007/s00146-020-01019-6, (2020.)

Kineska vlada snažno je poduprla nastojanja poduzeća da usvoje umjetnu inteligenciju, jasno vjerujući da ima ogroman potencijal za poticanje gospodarskog rasta. Njihov cilj je poticati industriju vrijednu 1 bilijun dolara i biti svjetski lider u umjetnoj inteligenciji do 2030.²⁷

Alibaba je jedna od vodećih b2b platformi za prodaju i suradnju. Milijuni proizvoda su dostupni na njihovim stranicama te predstavljaju velike napore za svakodnevno upravljanje njima. U tom dijelu dolaze u korist mogućnosti koje nudi umjetna inteligencija. Smanjenje troškova rada te povećana produktivnost i efikasnost procesa su ogromni.

Uvidjela smo ogromni poticaj kineske vlade za implementacijom ovog inovativnog sustava, štoviše Alibaba vješto koristi prednosti koje pruža. Kroz gustoću naseljenosti i vrstu vladavine koja zahvaća kinu, dostupnost i rukovanje podacima je pogodno za razvoj i korištenje mogućnosti umjetne inteligencije. Etičnost ovog zapleta je vrlo upitna.

„Alibabina ponuda zove se; platforma strojnog učenja za umjetnu inteligenciju, koja nudi rješenja za tvrtke koje žele iskoristiti prednosti kognitivnih računalnih funkcija kao što su obrada prirodnog jezika i računalni vid, bez početnih troškova izravnog ulaganja u infrastrukturu“²⁸.

Alibaba si postavlja velike ciljeve poput preoblikovanja gradova u „Pametne“ gradove, gdje bi se efikasnije upravljalo prometom i drugim funkcijama gradova. Nadalje jedan od ciljeva im je preoblikovati industriju agrokulture.

²⁷ CNBC: www.cnbc.com/2018/05/04/china-aims-to-steal-us-a-i-crown-and-not-even-trade-war-will-stop-it.html (23.04.2023.)

²⁸ Bernard Marr, Matt Ward, *Artificial Intelligence in Practice*, Wiley, 2019.

4. MEĐUZAVISNOST TRANSPARENTNOSTI, ETIKE I ODGOVORNOSTI

Manipulacija kroz umjetnu inteligenciju je realna prijetnja s kojom smo suočeni. Temeljni zahtjevi etičkog ponašanja ključni su za funkcionalnost i moral društva koje će biti sve više isprepletano s raznim funkcijama koje će imati umjetna inteligencija u našim svakidašnjim rutinama.

Transparentnost, etika i odgovornost čine jasnu predodžbu idealnog modela umjetne inteligencije, gdje se povjerenje u tehnologiju može stvoriti i održavati.

Daljnijim tekstom ulazi se dublje u temu kroz cjeline: 1.) Značajke vjerodostojnosti umjetne inteligencije, 2.) Glavne debate etike umjetne inteligencije.

4.1. Značajke vjerodostojnosti umjetne inteligencije

Interpretiranje i tumačenje te stvaranje vjerodostojnog i originalnog sadržaja, stvaraju veliki izazov kako bi sustav umjetne inteligencije mogao zadobiti povjerenje korisnika.

„Iako je možemo nazvati "pametnom", današnja umjetna inteligencija ne može sama razmišljati“²⁹.

U daljnjim cjelinama se dublje obrađuje vjerodostojnost sustava umjetne inteligencije.

4.1.1. Temeljni zahtjevi umjetne inteligencije

„Komponente umjetne inteligencije uključuju: strojno učenje, obradu prirodnog jezika, računalni vid, robotiku i ekspertne sustave. Ove komponente omogućuju strojevima da uče, razumiju i komuniciraju sa svijetom oko sebe na načine koji su prije bili nemogući“³⁰.

Navedeni su temeljni zahtjevi za definicijski funkcionalnu umjetnu inteligenciju. Potrebne komponente su ključ za stvaranje umjetne, ljudima-slične interakcije s ljudima.

4.1.2. Problematika vjerodostojnih sistema umjetne inteligencije na primjeru Michaela Millera

Primjer vjerodostojnosti umjetne inteligencije unutar industrije novinarstva, obraditi će se kroz primjer govora Michaela Millera (izvršni predsjednik u News Corp Australia), gosta na svjetskom kongresu novinskih medija 2023. godine u Taipeiu³¹.

²⁹ Caltech, *Can we trust artificial intelligence?:* www.scienceexchange.caltech.edu/topics/artificial-intelligence-research/trustworthy-ai (24.04.2023.)

³⁰ Dataconomy, *The building blocks of AI:* https://dataconomy.com/2023/04/03/basic-components-of-artificial-intelligence/?utm_content=cmp-true (22.05.2023.)

³¹ World Association of News Publishers, *AI companies crave credibility but it doesn't come for free:* www.wanfra.org/2023/04/ai-companies-crave-credibility-but-it-doesnt-come-for-free/ (26.04.2023.)

„Ponovno smo svjedoci dramatične promjene u digitalnom svijetu – a originalno novinarstvo i kreatori sadržaja moraju izbjeći greške iz prošlosti koje su desetkovale njihove industrije i omogućile tehnološkim tvrtkama da profitiraju, ali odbijaju platiti pravu vrijednost svoje kreativnosti“³².

Daljnje autor navodi kako razvoj umjetne inteligencije i munjevit promjene tehnologije i inovacije su unijele velike nestabilnosti na tržište. Tvrtke teško drže korak uz pomake i održavanje konkurentnosti je sve teže. Jedan od razloga povezan sa vjerodostojnosti je mogućnost sustava umjetne inteligencije da crpe tuđe ideje i kreativnosti kako bi stvorili svoj sadržaj. Takav proces je zasitio oglasne platforme te uveo određene promjene u potražnji.

„Generativna umjetna inteligencija, međutim, neće zamijeniti vještine novinara – da ispituju, istražuju i razvijaju svježije ideje i razmišljanja; objasniti kako pitanja i događaji mogu utjecati na budućnost; niti pružiti razumijevanje emocija, reakcija i iskustava koji oblikuju odluke u ljudskim životima ili oblikuju životnu sposobnost i ljubav zajednica“³³.

Navodi kako način na koji umjetna inteligencija prikuplja i stvara informacije je na prošlim događanjima te je trenutno ograničena svojim načinom rada. Kako bi se vjerodostojnost postigla, umjetna inteligencija bi trebala biti znatno ograničena i koristiti kao pomoćno sredstvo ili stvoriti svoj originalni sadržaj sa svježim idejama.

„Uz odgovarajuće zaštitne vodilje, umjetna inteligencija ima potencijal postati vrijedan novinarski resurs. Može pomoći novinarima i medijskim tvrtkama u brzom stvaranju sadržaja i prikupljanju činjenica“³⁴.

„Međutim, umjetna inteligencija otvara novi digitalni front, a kreatori sadržaja moraju učiti iz svoje prijašnje samodopadnosti. OpenAI je, primjerice, brzo uspostavio posao vrijedan 30 milijardi američkih dolara koristeći tuđi izvorni sadržaj i kreativnost bez naknade i atribucije“³⁵.

Slučaj vezan uz gore navedeni „OpenAI“ je detaljnije opisano „ChatGPT“, virtualni bot za razgovor koji omogućuje korisnicima da brzo dođu do odgovora i željenih informacija bez tradicionalnog samostalnog istraživanja Internet mreže. ChatGPT predstavlja tuđe informacije korisnicima te uzima zasluge za njih bez naknade izvorima tih informacija, ovim primjerom nam se javlja problem vjerodostojnosti.

„Jednostavno rečeno, AI sustavi suočavaju se s temeljnim rizikom za svoj budući uspjeh: uvjerenjem javnosti da su njihove informacije pouzdane i vjerodostojne, a da bi to postigli, morat će pošteno kompenzirati one koji pružaju sadržaj za njihov uspjeh“³⁶.

Dodatno na gore navedeni tekst, promjena u potražnji se dogodila iz razloga slabe vjerodostojnosti sustava umjetne inteligencije. Korisnici novinskih medija se sve više vraćaju kvalitetnim novinskim tvrtkama gdje su sigurni u kvalitetu i istinitost informacija na teško zasićenom oglasnom tržištu.

„Ovo može biti prava prilika za AI tvrtke, a ne prijetnja ili ručna kočnica. Partnerstva putem licenciranja i komercijalnih aranžmana s pouzdanim medijskim markama pružit će ono što im je tako očajnički potrebno za rast i uspjeh: vjerodostojnost“³⁷.

³² ibidem

³³ ibidem

³⁴ ibidem

³⁵ ibidem

³⁶ ibidem

³⁷ ibidem

4.2. Glavne debate etike umjetne inteligencije

Rizik koji predstavlja sustav umjetne inteligencije opisan je kroz debate, kako bi se optimiziralo plansko i strateško razdoblje.

U niže navedenim cjelinama proći ćemo kroz debate vezane uz etičko ponašanje sustava umjetne inteligencije i rizik koji donosi.

Niže navedene debate su obrađene na istraživanju Vincenta C. Mullera (Ethics of artificial intelligence and robotics)³⁸.

4.2.1. Privatnost i nadzor

Spektar digitalizacije se konstantno širi, naši život postaju sve više digitalizirani. Sva pohrana i prikupljanje podataka je digitalna te je odgovornost na sustavu velika. Implementacijom umjetne inteligencije stvara se mogućnost inteligentnog prikupljanja podataka i analizom istoga. Opći nadzor je iz tog razloga, vrlo dostižna mogućnost ukoliko se uzme za cilj. Razmjena tih podataka je poznata pojava među agentima, često u zamjenu za novac.

Kontrola i praćenje tko prikuplja koje podatke i tko ima pristup tim podacima je postalo gotovo nemoguće u digitalnom svijetu za razliku od analognog svijeta telefona i papira. Pojavom umjetne inteligencije, taj problem se pojačava s većinom današnjih modela. Kao primjer; prepoznavanje lica na fotografijama i videima dozvoljava identifikaciju, a time i profiliranje i traženje pojedinaca³⁹. Daljnje tehnologije sličnog karaktera koje ponekad budu otkrivene u „politici privatnosti“ rezultiraju; „U ovom golemom oceanu podataka postoji zastrašujuće potpuna slika o nama“⁴⁰. Rezultat nije dobio dužnu pozornost javnosti. Manipulira nas se da ostavljamo sve više podataka i da ostanemo nesvjesni o vrijednosti koju oni imaju. Za "velikih 5" tvrtki (Amazon, Google/Alphabet, Microsoft, Apple, Facebook) čini se da se glavni dio njihovog poslovanja koji se odnosi na prikupljanje podataka temelji na obmani, iskorištavanju ljudskih slabosti, promicanju odugovlačenja, stvaranju ovisnosti i manipulaciji⁴¹. „Nadzor je poslovni model interneta“⁴². Ovaj oblik ekonomije u kojoj su nadzor i pozornost ključni, naziva se „Nadzorni kapitalizam“⁴³. Čovjek

³⁸ Müller, Vincent C., "Ethics of Artificial Intelligence and Robotics", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N, 2021.

³⁹ Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Oscar Schwartz, Jason Schultz, AI Now Report 2018, 2018.

⁴⁰ Smolan, Sandy (director), 2016, "The Human Face of Big Data", *PBS Documentary*, 24 February 2016, 56 mins.

⁴¹ Harris, Tristan, *How Technology Is Hijacking Your Mind - from a Magician and Google Design Ethicist*, Thrive Global, 2016.

⁴² Schneier, Bruce, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, New York: W. W. Norton & Company, 2015.

⁴³ Zuboff, Shoshana, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: Public Affairs, 2019.

gubi vlasništvo nad svojim podacima te se može reći da se gubi kontrola nad našim podacima. Postoje tehnike i tehnologija za zaštitu podataka koji su danas standard očuvanje podataka i identiteta pojedinaca i grupa. Pojedine tvrtke su uvidjele prednosti za povećanje konkurentnosti putem zaštite podataka te se može iskoristiti putem prodaje i korištenjem „poluga“ u razmjeni.

Provedba regulativnih mjera na razini države i pojedinca je jedan od velikih izazova povezanih uz digitalni svijet. Sud još nije dovoljno kompetentan kako bi proveo suđenje i tumačio problematiku i njenu rješenje. Zaštita dobro ustanovljenih prava još nije čvrstog temelja u digitalnim proizvodima te se koristi od strane poduzeća kao mogućnost testiranja svojih proizvoda na korisnicima bez straha ili odgovornosti, nadalje braneći svoja prava intelektualnog vlasništva.

4.2.2. Manipulacija ponašanja

Problemi etike umjetne inteligencije nadilaze površno sakupljanje podataka. Podaci se koriste na način da manipuliraju ponašanje, na mreži i izvan nje. Na način koji potkopava čovjekov autonomni racionalni odabir. Tehnike manipulacije ponašanja postoje još od pamtivijeka, pojavom umjetne inteligencije, taj problem bi mogao steći novi glas. Korisnici su u tolikoj mjeri ovisni te u interakciji sa sustavima podatak i znanjem pruženim kroz mrežu, da su skloni i ranjivi prema manipulaciji, prijevari i „Gurkanju“. Uz dovoljnu količinu podataka, sustavi umjetne inteligencije se mogu koristiti na način da ciljaju određenu skupinu pojedinaca sa određenim karakteristikama te utjecali na njih sa izabranim informacijama koje efektivno utječu na njih. „Gurkaj“ mijenja okolinu tako da utječe na ponašanje na predvidljiv način koji je pozitivan za pojedinca, ali ga je lako i jeftino izbjeći⁴⁴. Nijanse dijele trenutno stanje od paternalizma i manipulacije.

Mnogi oglašivači, trgovci i online prodavači koriste će sva pravna sredstva koja su im na raspolaganju kako bi povećali profit, uključujući iskorištavanje predrasuda u ponašanju, prijevaru i stvaranje ovisnosti⁴⁵. U dizajnu sučelja na web stranicama ili u igrama, ova manipulacija koristi ono što se naziva "tamnim uzorcima"⁴⁶. Stroga regulacija se provodi u poslovanju kockanja i opojnih sredstava, no u digitalnom svijetu ne.

Socijalne mreže postaju glavna infrastruktura za političku propagandu i utjecaj na glasanje. Štoviše, lažiranje umjetne inteligencije pretvara što su nekad bili čvrsti dokazi, u nepouzdanje i samim time se oslabljuju unošenje temelja za regulacije. U isto vrijeme će nepouzdanje i nepovjerenje u digitalne interakcije rasti, dok s druge strane smo sve više ovisni o njima kako bi tijekom naših promijenjenih života tekao zamišljenom crtom.

Područje politike vezano uz ovu temu je pod velikim pritiskom od strane poslovnih lobista, tajnih službi i agencija koje ovise o nadzornim aktivnostima. Zaštita podataka je skoro pa nestala u usporedbi sa prijašnjim vremenima gdje je komunikacija bila temeljena na pismima, analognim telefonima i osobnom komunikacijom te kada je nadzor funkcionirao pod značajnim restrikcijama.

⁴⁴ Thaler, Richard H., Sunstein, Cass, *Nudge: Improving decisions about health, wealth and happiness*, New York: Penguin, 2008.

⁴⁵ Costa, Elisabeth and David Halpern, "The Behavioural Science of Online Harm and Manipulation, and What to Do About It: An Exploratory Paper to Spark Ideas and Debate", The Behavioural Insights Team Report, 1-82, 2019.

⁴⁶ Mathur, Arunesh, Gunes Acar, Michael J. Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan, "Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites", *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW): art. 81. doi:10.1145/3359183, 2019.

Opća uredba EU o zaštiti podataka⁴⁷ ojačala je zaštitu privatnosti, dok SAD i Kina preferiraju rast uz slabiju regulaciju⁴⁸.

4.2.3. Autonomija sistema

Autonomija sistema ne implicira na odgovornost, dok suprotno; odgovornost implicira na autonomiju. Drugim riječima, postoje sistemi koji imaju određeni stupanj tehničke autonomije, bez uzrokovanja problematike odgovornosti. Slabiji, više tehnički, pojam autonomije u robotici je relativan i postupan: kaže se da je sustav autonoman u odnosu na ljudsku kontrolu do određenog stupnja.

Do kojeg stupnja je autonomija sustava funkcionalna bez stvaranja raznih problema? Pitanje se svodi na određivanje rješavanja te problematike kroz adaptacije kontekstualnog okvira ili tehničko podešavanje. U većini pravosuđa postoji odjel koji razrješava takva pitanja povezana uz problematičnost umjetne inteligencije. Dok će se npr. U određenim sustavima će se tehnički standardi morati izmjenjivati s ciljem implementacije novih tehnologija poput umjetne inteligencije.

Tijela poput IEEE-a (Institut inženjera elektronike i elektrotehnike) i BSI (Britanski institut za standarde) stvaraju standarde za problematiku vezanu uz debate etike umjetne inteligencije.

Ponekad se postavljaju pitanja; kada bi sistemi trebali striktno pratiti pravila, a kada stvoriti kompromis radi veće koristi.

4.2.4. Prozirnost

Neprozirnost i pristranost središnja su pitanja u onome što se sada ponekad naziva "etikom podataka" ili "etikom velikih podataka"⁴⁹. Sustavi umjetne inteligencije za automatiziranu podršku odlučivanju i "predvidivu analitiku" izazivaju "značajnu zabrinutost zbog nedostatka odgovarajućeg postupka, odgovornosti, angažmana zajednice i revizije"⁵⁰. Pogođena osoba neće znati kako je sustav došao do rezultat, sustav je „neproziran“ za tu osobu. Ukoliko sustav sam uči kroz takozvano „strojno učenje“, rezultat će i stručnjaku biti nepoznanica te neće znati kako je nastao uzorak stvoren od strane sustava. Ova neprozirnost pogoršava pristranost u skupovima podataka i sustavima odlučivanja.

Politička intervencija mora obraditi pristranost i neprozirnost zajedno, kako bi došli do zdravog zaključka.

⁴⁷ Regulation (EU) 2016/679, 2016.

⁴⁸ Thompson, Nicholas and Ian Bremmer, 2018, "The AI Cold War That Threatens Us All", *Wired*, 2018.

⁴⁹ Mittelstadt, Brent Daniel and Luciano Floridi, "The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts", *Science and Engineering Ethics*, 22(2): 303–341. doi:10.1007/s11948-015-9652-2, 2016.

Floridi, Luciano and Mariarosaria Taddeo, "What Is Data Ethics?", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083): 20160360. doi:10.1098/rsta.2016.0360, 2016.

⁵⁰ Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Oscar Schwartz, Jason Schultz, AI Now Report 2018, 2018.

Većina sustava umjetne inteligencije se bazira na strojnom učenju te stvara uzorak iz skupa podataka bez potrebe da je točno rješenje. Stručnjak ne može biti siguran u rezultat i način operativnosti sustava. Sustavu se može dati do znanja gdje je rješenje bilo točno, a gdje netočno. Time bi se sustav mogao prilagoditi i adaptirati način operacionalizacije. Ništa nije transparentno, već je neprozirno. Kvaliteta podataka dana sustavu je ključna za stvaranje kvalitetnih rezultata.

U Europi je ova problematika uzeta u obzir unutar svoje (Regulation (EU) 2016/679). Vezano uz neprozirnost podataka; korisnik će imati legalnu mogućnost za objašnjenje rezultata stvorenog kroz procesiranje podataka uz pomoć umjetne inteligencije.

4.2.5. Singularnost

Ljudska inteligencija, kao i svaka druga inteligencija bilo koje druge biološke vrste, jest dispoziციjski nadarena za sustavnu obradu neophodnu za proizvodnju prosudba, u skladu sa svojom prirodom, u sučelju preživljavanja⁵¹.

Postizanjem ljudske razine inteligencije u sustavima umjetne inteligencije, postiže se razina operativnosti u kojoj je moguće da umjetna inteligencija stvori svoju vrstu umjetne inteligencije koja bi posvojila pojam „super-inteligencije“. Stanje u kojemu stvaranje i kontroliranje sistema izmiče ljudskoj kontroli.

Strahovi o robotima ustanovljeni još prije postojanja umjetne inteligencije⁵², pokazali su se stvarnim i mogućim. Eksplozivni rast i razvoj tehnologije razotkriva mogućnosti koje su se teško mogle predvidjeti u prijašnjim slabije razvijenim uvjetima.

Pojam singularnost predstavlja hipotetsku budućnost u kojoj razvoj i rast umjetne inteligencije postaje nekontroliran i nepovratan. Realnost bi se mogla radikalno i nepredvidivo promijeniti.

Znanstvenici i teoretičari pokušavaju odrediti vrijeme i trenutak kada bi se postigla singularnost. Jedan od mjernih subjekata je vještina koju umjetna inteligencija poprima nalik čovjeku, poput jezika. Jezik je najizazovnija prepreka za umjetnu inteligenciju te bi model umjetne inteligencije koji suzi odstupanje bio najbliža točka prema singularnosti u ovome trenutku.

Budućnost koja nas očekuje je nepredvidiva, istina će se saznati kroz vrijeme. Singularnost bi se mogla postići već ovo desetljeće po mišljenju stručnjaka, dok neki tvrde da još nije ni blizu. Sve ovisi o razvoju umjetne inteligencije. Oprezne regulacije su ključne za držanje stanja pod kontrolom, neoprezno davanje pristupa umjetnoj inteligenciji donosi kobne posljedice koji će svakako rezultirati nepoželjnim te vjerojatno nepovratnim greškama.

⁵¹ Maria Odete Madeira & Carlos Pedro Gonçalves, *AI Awakening – A Technological Singularity*, 2017.

⁵² Samuel Butler, *Darwin Among the Machines*, The Press, 1863.

4.2.6. Automatizacija i zapošljavanje

Produktivnost i bogatstvo će se kroz umjetnu inteligenciju znatno povećati. Automatizacijom se pospješuje proces proizvodnje te je potrebno manje ljudi za izvršenje tog procesa. Ne mora nužno značiti da će bit manjak radnih mjesta, no suprotno od toga, više bogatstva stvara veću potražnju koja može negirati rast produktivnosti.

U 20 godina između 1950. i 1970. broj unajmljenih poljoprivrednih radnika u Velikoj Britaniji smanjen je za 50%⁵³. Rezultat toga je preseljenje industrija u mjesta s jeftinijim uvjetima rada.

Klasična automatizacija zamijenila je ljudske mišice, dok digitalna automatizacija zamjenjuje ljudsku misao ili obradu informacija—a za razliku od fizičkih strojeva, digitalna automatizacija vrlo je jeftina za dupliciranje⁵⁴. Jeftino dupliciranje može rezultirati radikalnijom promjenom radnog tržišta. Teško se prognozira ako će stvaranje novih radnih mjesta nadomjestiti ona izgubljena. Nadalje, hoće li trošak tranzicije i adaptacije radne snage za nova radna mjesta biti znatan? Tko snosi troškove, a tko grabi korist?

Možda će enormna povećanja produktivnosti omogućiti da se ostvari "doba dokolice"⁵⁵. Harari objašnjava kako je ovaj ekonomski razvoj omogućio čovječanstvu da prevlada glad, bolest i rat — a sada ciljamo na besmrtnost i vječno blaženstvo putem umjetne inteligencije, stoga njegov naslov „Homo Deus“⁵⁶.

Trenutno stanje je pod prijetnjom nepravedne distribucije bogatstva kroz slobodna tržišta gdje „pobjednik uzima sve“. Trenutno postoje naznake takvog razvoja.

Zanimljivo pitanje je ako je umjetna inteligencija održiva; umjetna inteligencija proizvodi puno otpada kao i ostali kompjutorski sistemi te se ti troškovi čak i otplaćuju putem društva.

4.2.7. Pristranost odlučivanja

Autor navodi kako sistemi umjetne inteligencije služe kao alati za rješavanje potreba ljudi. Brže i vještije izračune te sveukupno predstavljaju nadogradnju na određeno polje primjene. Služe kao način predviđanja budućih ishoda te se primjenjuju u poslovnim, zdravstvenim i ostalim područjima. Takva mogućnost predviđanja je jeftinija solucija od dosadašnjih.

Nadalje, sustav je podložan pristranosti odlučivanja kroz određivanje karakteristika subjekta nad kojim se vrši odluka. Unosom podataka i često nerelevantnih karakteristika subjekta, sustav odlučivanja donosi preventivne odluke, štoviše ne dozvoljava činu da dosegne krajnji ishod kako bi

⁵³ Zayed, Yago and Philip Loft, 2019, "Agriculture: Historical Statistics", *House of Commons Briefing Paper*, 3339, 1-19., 2019.

⁵⁴ Bostrom, Nick and Eliezer Yudkowsky, "The Ethics of Artificial Intelligence", in *The Cambridge Handbook of Artificial Intelligence*, Keith Frankish and William M. Ramsey (eds.), Cambridge: Cambridge University Press, 316–334. doi:10.1017/CBO9781139046855.020, 2014.

⁵⁵ Keynes, John Maynard, "Economic Possibilities for Our Grandchildren". Reprinted in his *Essays in Persuasion*, New York: Harcourt Brace, 1932, 358–373, 1930.

⁵⁶ Harari, Yuval Noah, *Homo Deus: A Brief History of Tomorrow*, New York: Harper, 2016.

„sigurnost“ bila što veća. Upitnost sigurnosti je diskutabilna zbog nerelevantnih i nekvalitetnih podataka o subjektu zbog kojih se stvara sivo područje pogodno manipulaciji. Pristranost je moguće uvijek povećati već u vrsti podataka koji su temelj trenutnog sustava (npr. Povećanom aktivnošću rudarenja, otkrije se veća količina zlata u određenom mjestu).

Povjerenje u tehničku kvalitetu sustava i ciljeve određenog sektora u kojem je implementirana, moglo bi se utvrditi ukoliko je problematika postojeća.

Nastavlja kako pristranost se često pojavljuje kod nepravedne osude, kada se pojedinac nepravedno osudi po karakteristikama „uže“ grupe ljudi u koju je svrstan te nije sveden na jednostrane odlike. Sustav krivim tumačenjem i nepostojećim opsežnim razmatranjem situacije, vrlo jednostavno je postići nekvalitetnu osudu prema načelima pravednosti. Što je zapravo pravednost? Znači li pravednost ispunjenje načela i ideala prema neutralnom i ujednačenom sustavu etičkih i moralnih vrijednosti, iako su ljudi po prirodi skloni pristranosti. Kognitivna sposobnost ljudi za sveobuhvatno tumačenje je omogućila evolucijski rast uz prisustvo pristranosti, nadalje važna je vještina koju bi umjetna inteligencija trebala imati u određenom obliku za donošenje nama ljudima pravednih odluka. Postavlja se pitanje ako bi umjetna inteligencija uopće trebala imati određene ljudske vrline poput intuitivne pristranosti te ako je to moguće.

Postoje značajni tehnički naponi da se otkrije i ukloni pristranost iz AI sustava, ali pošteno je reći da su oni u ranim fazama; Britanski institut za etičku umjetnu inteligenciju i strojno učenje⁵⁷.

Autor govori kako set podataka korišten za stvaranje nepristrane odluke, ne bi se trebao ponovno koristiti. Razlog tome je stvaranje uzorka i povijesti vezane uz pristranost unutar tih podataka koje čovjek zna tumačiti za razliku od umjetne inteligencije. Primjer Amazonovog automatiziranog sustava provjere zapošljavanja, gdje je otkrivena vjerojatnost diskriminacije nad zapošljavanjem žena (prekinuto je 2017. godine). Određeni set podataka u mogućnosti je stvoriti, ne više od jedne nepristrane odluke te iz tog razloga postoji opasnost od krivog sekundarnog tumačenja u novom scenariju.

„Dokazi sugeriraju da modeli umjetne inteligencije mogu ugraditi i implementirati ljudske i društvene predrasude na razmjeru. Međutim, temeljni podaci su odgovorni, a ne sam algoritam“⁵⁸.

⁵⁷ Brownsword, Roger, Eloise Scotford, and Karen Yeung (eds.), 2017, *The Oxford Handbook of Law, Regulation and Technology*, Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199680832.001.0001; Yeung, Karen and Martin Lodge (eds.), 2019, *Algorithmic Regulation*, Oxford: Oxford University Press. doi:10.1093/oso/9780198838494.001.0001

⁵⁸ Gregory S. Nelson, *Bias in Artificial Intelligence*, PMID: 31278182, DOI:10.18043/ncm.80.4.220, 2019

5. ZAKLJUČAK

Promatrajući trenutni tijek implementacije umjetne inteligencije u određene poslovne grane, svjedočimo velikim uspjesima i u istom trenutku velikim sumnjama i propitkivanjima vezanim uz način funkcioniranja i etičnog postupanja sustava. Teorije, smjernice i preporuke stručnjaka i poslovnih ljudi su u načelu pismeno možda vjerodostojne, ali u praktičnoj implementaciji gube moralnu vrijednost iz teorijske faze planiranja. Razlozi tomu su nedokazani i skloni debatiranju i svakojakim mišljenjima, no svakako je vrijedno reći da su određeni moralno upitni interesi s najviše utjecaja unutar spektra krivice. Teorijskim razlaganjem smjernica i stvaranjem preporuka za implementaciju i korištenje sustava umjetne inteligencije, inicira se stvoriti što veći ekonomski koristan učinak uz što manje negativnih posljedica. Potreba za stvaranjem konkurentnosti i opstankom unutar ekonomske utrke stvara strašan pritisak na konkurentne zemlje i poduzeća unutar nje. Mogućnosti pogreške i preranih riskantnih odluka zbog ekonomske utrke su neporecivo prisutne i sklone nanijeti veliku štetu.

Etičnost unutar zdravstva poslužio nam je kao dobar primjer važnosti određenih etičnih smjernica kao što su pristranost, manipulacija i privatnost podataka. Važnost ovih etičnih načela jedni su od današnjih temeljnih zahtjeva za prava pojedinca. Uvidjeli smo komplikaciju koja se pojavljuje prilikom razvitka i nastojanja osposobljavanja sustava za uspješno provođenje procesa po zadanim moralnim načelima. Kompleksnost sustava i neslaganje određenih moralnih zahtjeva i načina funkcioniranja trenutnog sustava dovode zastoje i upitni tijek razvoja nekih moralnih načela. Određene zemlje s visoko digitaliziranim zdravstvenim sustavom su jedine koje će u bližoj budućnosti imati opciju za implementaciju umjetne inteligencije. Sklonost upitnim odlukama sustava nakon implementacije umjetne inteligencije je među glavnim problemima ako se govori o sustavu visokog stupnja autonomnosti. Nemir među pacijentima popratni je ishod prije navedene problematike te se pitanje blagostanja i mira unutar društva narušava.

Zaključno možemo reći kako je inovativna tehnologija etične umjetne inteligencije u svojim prvim koracima, implementacija, funkcionalnost i primjena su još u pitanju te će u bližoj budućnosti možda biti uspješno implementirana u poslovne sustave. Trenutna korist umjetne inteligencije se svodi na ulogu korisnog alata koji ima pojedine mogućnosti i sposobnosti funkcionalnosti nalik čovjekovoj. Prednosti mogu biti velike i revolucijskog karaktera, donoseći promjenu u svakidašnji život pojedinca i poduzeća, dok iz druge perspektive nanosi pritisak i nužnost na promjenu i adaptaciju kroz koju mogu nastati interne i eksterne nesuglasice i nestabilnosti.

LITERATURA

INTERNET IZVORI

1. C3.ai, *Glossary, Ethical AI*: www.c3.ai/glossary/artificial-intelligence/ethical-ai/ (20.04.2023.)
2. Organizacija za gospodarsku suradnju i razvoj, *OECD.ai, Investiranje u istraživanje i razvoj umjetne inteligencije (Načelo 2.1)*: <https://oecd.ai/en/dashboards/ai-principles/P10> (22.04.2023.)
3. Organizacija za gospodarsku suradnju i razvoj, *OECD.ai, Poticanje digitalnog ekosustava za umjetnu inteligenciju (Načelo 2.2)*: <https://oecd.ai/en/dashboards/ai-principles/P11> (22.04.2023.)
4. Organizacija za gospodarsku suradnju i razvoj, *OECD.ai, Oblikovanje poticajnog političkog okruženja za umjetnu inteligenciju (Načelo 2.3)*: <https://oecd.ai/en/dashboards/ai-principles/P12> (22.04.2023.)
5. Organizacija za gospodarsku suradnju i razvoj, *OECD.ai, Izgradnja ljudskih kapaciteta i priprema za transformaciju tržišta rada (Načelo 2.4)*: <https://oecd.ai/en/dashboards/ai-principles/P13> (22.04.2023.)
6. Organizacija za gospodarsku suradnju i razvoj, *OECD.ai, Međunarodna suradnja za pouzdanu umjetnu inteligenciju (Načelo 2.5)*: <https://oecd.ai/en/dashboards/ai-principles/P14> (22.04.2023.)
7. The White House, *Blueprint for AI Bill of Rights*: www.whitehouse.gov/ostp/ai-bill-of-rights/ (23.04.2023.)
8. Europska komisija, *Javna uprava i zaštita podataka, Koji su glavni aspekti Opće uredbe o zaštiti podataka (GDPR) kojih bi javna uprava trebala biti svjesna?*: https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/public-administrations-and-data-protection/what-are-main-aspects-general-data-protection-regulation-gdpr-public-administration-should-be-aware_en#references (23.04.2023.)
9. Microsoft, *Responsible and trusted AI*: <https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/innovate/best-practices/trusted-ai> (23.04.2023.)
10. Businesswire, *News*: www.businesswire.com/news/home/20230131005091/en/73-of-Consumers-Believe-AI-can-have-a-Positive-Impact-on-their-Customer-Experience (23.04.2023.)
11. CNBC: www.cnbc.com/2018/05/04/china-aims-to-steal-us-a-i-crown-and-not-even-trade-war-will-stop-it.html (23.04.2023.)
12. Caltech, *Can we trust artificial intelligence?*: www.scienceexchange.caltech.edu/topics/artificial-intelligence-research/trustworthy-ai (24.04.2023.)
13. Dataconomy, *The building blocks of AI*: https://dataconomy.com/2023/04/03/basic-components-of-artificial-intelligence/?utm_content=cmp-true (22.05.2023.)
14. World Association of News Publishers, *AI companies crave credibility but it doesn't come for free*: www.wan-ifa.org/2023/04/ai-companies-crave-credibility-but-it-doesnt-come-for-free/ (26.04.2023.)
15. TechTarget, *WhatIs, AI Ethics (AI code of ethics)*: <https://www.techtarget.com/whatis/definition/AI-code-of-ethics> (27.04.2023.)
16. T. Dušić, "Etika umjetne inteligencije", Diplomski rad, Sveučilište u Rijeci, Filozofski fakultet, Rijeka, 2021. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:186:018413> (28.05.2023.)
17. Ž. Puzek, "Etika umjetne inteligencije", Diplomski rad, Sveučilište u Zagrebu, Filozofski fakultet, Zagreb, 2018. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:131:795477> (28.05.2023.)
18. V. Jaklić, "Utjecaj umjetne inteligencije u poslovanju", Diplomski rad, Sveučilište u Zagrebu, Ekonomski fakultet, Zagreb, 2019. Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:148:363157> (28.05.2023.)

KNJIGE

1. Schneier, Bruce, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, New York: W. W. Norton & Company, 2015.
2. Zuboff, Shoshana, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, New York: Public Affairs, 2019.
3. Harris, Tristan, *How Technology Is Hijacking Your Mind - from a Magician and Google Design Ethicist*, Thrive Global, 2016.
4. Thaler, Richard H., Sunstein, Cass, *Nudge: Improving decisions about health, wealth and happiness*, New York: Penguin, 2008.
5. Keynes, John Maynard, "Economic Possibilities for Our Grandchildren". Reprinted in his *Essays in Persuasion*, New York: Harcourt Brace, 1932, 358–373, 1930.
6. Samuel Butler, *Darwin Among the Machines*, The Press, 1863.
7. Harari, Yuval Noah, *Homo Deus: A Brief History of Tomorrow*, New York: Harper, 2016.
8. Brownsword, Roger, Eloise Scotford, and Karen Yeung (eds.), *The Oxford Handbook of Law, Regulation and Technology*, Oxford: Oxford University Press.
doi:10.1093/oxfordhb/9780199680832.001.0001, 2017.
9. Gregory S. Nelson, *Bias in Artificial Intelligence*, PMID: 31278182, DOI:10.18043/nem.80.4.220, 2019.
10. Bernard Marr, Matt Ward, *Artificial Intelligence in Practice*, Wiley, 2019.

ČLANCI

1. Helen Smith, Clinical AI: opacity, accountability, responsibility and liability. (2020) 36:535–45. doi: 10.1007/s00146-020-01019-6, (2020.)
2. Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Oscar Schwartz, Jason Schultz, AI Now Report 2018, 2018.
3. Maria Odete Madeira & Carlos Pedro Gonçalves, AI Awakening – A Technological Singularity, 2017.
4. Regulation (EU) 2016/679, 2016.
5. Costa, Elisabeth and David Halpern, "The Behavioural Science of Online Harm and Manipulation, and What to Do About It: An Exploratory Paper to Spark Ideas and Debate", The Behavioural Insights Team Report, 1-82, 2019.
6. Mathur, Arunesh, Gunes Acar, Michael J. Friedman, Elena Lucherini, Jonathan Mayer, Marshini Chetty, and Arvind Narayanan, "Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites", *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW): art. 81. doi:10.1145/3359183, 2019.
7. Müller, Vincent C., "Ethics of Artificial Intelligence and Robotics", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N, 2021.
8. Thompson, Nicholas and Ian Bremmer, 2018, "The AI Cold War That Threatens Us All", *Wired*, 2018.
9. UNESCO, Recommendations on the Ethics of Artificial Intelligence, SHS/BIO/REC-AIETHICS/2021, 2021.

10. Mittelstadt, Brent Daniel and Luciano Floridi, “The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts”, *Science and Engineering Ethics*, 22(2): 303–341. doi:10.1007/s11948-015-9652-2, 2016.
11. Floridi, Luciano and Mariarosaria Taddeo, “What Is Data Ethics?”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083): 20160360. doi:10.1098/rsta.2016.0360, 2016.
12. Zayed, Yago and Philip Loft, 2019, “Agriculture: Historical Statistics”, *House of Commons Briefing Paper*, 3339, 1-19., 2019.
13. Bostrom, Nick and Eliezer Yudkowsky, “The Ethics of Artificial Intelligence”, in *The Cambridge Handbook of Artificial Intelligence*, Keith Frankish and William M. Ramsey (eds.), Cambridge: Cambridge University Press, 316–334. doi:10.1017/CBO9781139046855.020, 2014.
14. Yeung, Karen and Martin Lodge (eds.), *Algorithmic Regulation*, Oxford: Oxford University Press. doi:10.1093/oso/9780198838494.001.0001, 2019.
15. Brownsword, Roger, Eloise Scotford, and Karen Yeung (eds.), *The Oxford Handbook of Law, Regulation and Technology*, Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199680832.001.0001, 2017.

DOKUMENTANI FILMOVI

1. Smolan, Sandy (director), 2016, “The Human Face of Big Data”, *PBS Documentary*, 56 minuta, 2016.