

# Regulatorni okvir za oprabu umjetne inteligencije u poslovanju

---

**Družinec, Marin**

**Master's thesis / Diplomski rad**

**2024**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Economics and Business / Sveučilište u Zagrebu, Ekonomski fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:148:491643>

*Rights / Prava:* [Attribution-NonCommercial-ShareAlike 3.0 Unported/Imenovanje-Nekomercijalno-Dijeli pod istim uvjetima 3.0](#)

*Download date / Datum preuzimanja:* **2025-01-28**



*Repository / Repozitorij:*

[REPEFZG - Digital Repository - Faculty of Economics & Business Zagreb](#)



**Sveučilište u Zagrebu**  
**Ekonomski fakultet**  
**Integrirani preddiplomski i diplomski sveučilišni studij**  
**Poslovna ekonomija – smjer Trgovinsko poslovanje**

**REGULATORNI OKVIR ZA UPORABU UMJETNE INTELIGENCIJE**  
**U POSLOVANJU**

**Završni rad**

**Marin Družinec**

**Zagreb, 2024.**

**Sveučilište u Zagrebu**  
**Ekonomski fakultet**  
**Integrirani preddiplomski i diplomski sveučilišni studij**  
**Poslovna ekonomija – smjer Marketing**

**REGULATORNI OKVIR ZA UPORABU UMJETNE INTELIGENCIJE**  
**U POSLOVANJU**

**REGULATORY FRAMEWORK FOR THE USE OF ARTIFICIAL**  
**INTELLIGENCE IN BUSINESS**

**Završni rad**

**Student: Marin Družinec**  
**Mentor: Zvonimir Šafranko**

**Zagreb, 2024.**

# Sadržaj

1. UVOD .....	1
1.1. Predmet i cilj rada.....	1
1.2. Istraživačka pitanja .....	2
2. TEORIJSKO DEFINIRANJE I KATEGORIZACIJA UMJETNE INTELIGENCIJE .....	3
2.1. Definiranje koncepta umjetne inteligencije .....	3
2.2. Tipovi umjetne inteligencije .....	6
2.3. Utjecaj umjetne inteligencije na poslovne procese.....	10
2.4. Primjeri upotrebe umjetne inteligencije .....	11
2.5. Budući trendovi i izazovi u razvoju umjetne inteligencije .....	14
3. REGULATORNI OKVIR ZA UPORABU UMJETNE INTELIGENCIJE U POSLOVANJU - AKT O UMJETNOJ INTELIGENCIJI .....	17
3.1. Pravne osnove, načelo supsidijarnosti i proporcionalnosti.....	17
3.2. Zabranjeni AI sustavi i visoko rizični AI sustavi .....	19
3.3. AI opće namjene (GPAI).....	25
3.4. Upravljanje i provedba EU Akta .....	28
4. ZAKLJUČAK .....	30
POPIS LITERATURE .....	31

# 1. UVOD

Prijedlog Akta o umjetnoj inteligenciji (EU AI Act) razvio se kao odgovor na sve veći utjecaj AI tehnologija na svakodnevni život građana Europske unije. Ovaj zakonodavni okvir prvi je put predstavljen u travnju 2021. godine, a u tijeku su rasprave i pregovori između Europskog parlamenta, Vijeća EU-a i Europske komisije kako bi se usvojile konačne odredbe. Dok se prijedlog razvija, u fokusu su osiguravanje sigurnosti, zaštita ljudskih prava i poticanje inovacija na jedinstvenom digitalnom tržištu EU-a.

Akt o umjetnoj inteligenciji predstavlja regulatorni okvir dizajniran da upravlja rizicima i prilikama povezanim s primjenom umjetne inteligencije (AI) unutar Europske unije. Ovaj zakonodavni prijedlog, koji je prvi te vrste na svijetu, ima za cilj uspostaviti standarde za pouzdanu i sigurnu upotrebu AI tehnologija, čime se osigurava da inovacije doprinose društvenom i gospodarskom razvoju, a istovremeno štite prava građana EU. Ključni aspekti Akta uključuju stroge zahtjeve za visokorizične aplikacije AI-a, kao što su one koje se koriste u zdravstvenoj zaštiti, policijskim nadzorima i zapošljavanju, gdje greške ili zlouporabe mogu imati ozbiljne posljedice za pojedince.

Akt također promiče transparentnost i odgovornost u razvoju AI tehnologija, zahtijevajući od tvrtki da dokumentiraju kako njihovi AI sustavi rade i da objasne odluke koje ti sustavi donose. To uključuje obvezu objavljivanja izvora podataka koji se koriste za treniranje AI modela, što je ključno za sprječavanje pristranosti i osiguranje pravičnosti. S ovim regulativama, EU ne samo da nastoji zaštititi svoje građane, već i potaknuti globalnu raspravu i postavljanje standarda u razvoju umjetne inteligencije, čime se pozicionira kao globalni lider u etičkom pristupu tehnološkim inovacijama.

## 1.1. Predmet i cilj rada

Predmet rada je detaljna analiza Aktu EU o umjetnoj inteligenciji, sa posebnim fokusom na njegovu strukturu, ključne odredbe, i potencijalne učinke na različite sektore unutar Europske unije. Rad bi također mogao istraživati kako se ovaj regulativni okvir uklapa u širi kontekst globalnih normi i standarda u području umjetne inteligencije, te kakve su njegove implikacije za međunarodnu suradnju i konkurentnost u tehnološkom sektoru.

Cilj rada je procijeniti učinkovitost Aktu EU o umjetnoj inteligenciji u postizanju svojih temeljnih ciljeva, kao što su zaštita prava pojedinaca i promicanje etičkog razvoja i primjene AI tehnologija. Rad teži identificirati i raspraviti ključne izazove i mogućnosti koje Akt predstavlja za poduzeća, istraživačke institucije i policy-makers. Osim toga, rad bi mogao pružiti preporuke za poboljšanje trenutnog zakonodavnog okvira, na temelju analize prikupljenih podataka i studija slučaja. Cilj je također povećati razumijevanje utjecaja ovog značajnog zakonodavnog dokumenta među zainteresiranim stranama i širom javnosti.

## 1.2. Istraživačka pitanja

U skladu s predmetom i ciljem rada postavljaju se istraživačka pitanja:

1. Kako Akt EU o umjetnoj inteligenciji utječe na konkurentnost europskih tehnoloških poduzeća?
2. U kojoj mjeri su odredbe Akta o umjetnoj inteligenciji uspješne u minimiziranju rizika povezanih s pristranošću i diskriminacijom u AI aplikacijama?
3. Kakve izazove i prepreke susreću implementatori i regulatori u provođenju Aktu EU o umjetnoj inteligenciji?

## 2. TEORIJSKO DEFINIRANJE I KATEGORIZACIJA UMJETNE INTELIGENCIJE

### 2.1. Definiranje koncepta umjetne inteligencije

Umjetna inteligencija (AI) širok je skup računalnih sustava koji imaju sposobnost samostalnog učenja i donošenja odluka. Od početka razvoja umjetne inteligencije, simulacija inteligentnih ponašanja koja su očigledna ljudskim bićima uvijek je ostala temeljni kapacitet tehnologije. Različite varijante Turingovog testa osmišljene su kako bi se ispitalo može li stroj pokazati inteligenciju koja se ne može razlikovati od ljudi, ili čak bolju, inteligenciju. Suvremena definicija umjetne inteligencije često obuhvaća ove osnovne kriterije, ali također inzistira na uključivanju sposobnosti samostalnog djelovanja i specifičnih inteligencija koje se koriste u mnogim domenama (Boucher, 2020).

U međuvremenu su čak razvijene aplikacije AI tehnologija za usavršavanje replikacija misaonih procesa i donošenja odluka kod ljudi, ali ih dodatno optimiziraju za posebne primjene u medicinskoj dijagnostici, autonomnoj vožnji vozila ili kontroli teških industrijskih procesa . U tom kontekstu, AI sustavi dizajnirani su za prepoznavanje obrazaca, donošenje odluka i djelovanje na temelju velikih količina podataka uz visok stupanj neovisnosti o ljudskom nadzoru (Boucher, 2020).

Umjetna inteligencija (AI) sposobnost je računalnog sustava da oponaša ljudsku percepciju, učenje, donošenje odluka i obradu jezika. U temelju te raznolikosti i inkluzivnosti nalazimo prividnu inteligenciju kojoj gravitira mnoštvo interpretacija. Suvremene rasprave o umjetnoj inteligenciji često se vrte oko mnoštva pitanja, od toga koliko će biti nalik čovjeku do postojećih AI sustava, onih koji su trenutno u razvoju i onih iz znanstvene fantastike. Širina primjene otežava preciznu procjenu mogućih učinaka koje bi razvoj umjetne inteligencije mogao imati,

budući da bi upotrebe i rizici mogli biti vrlo različiti u različitim kontekstima u kojima se AI tehnologije implementiraju (Haenlein i sur., 2019).

Da bi se rasprave o AI-u učinile smislenijim i konstruktivnijim, potrebno je preciznije razlikovati među različitim vrstama AI sustava i kontekstima u kojima se koriste. Razlike između jednostavnih sustava koji služe savjetodavnim ulogama i složenih algoritama koji samostalno donose odluke su značajne, osobito kada ti algoritmi utječu na pojedinačne živote. Također je ključno razlikovati između hipotetičkih budućih razvoja koji možda nikada neće biti realizirani i tehnologija koje već danas imaju konkretan utjecaj na društvo.

Definicije umjetne inteligencije su se tijekom vremena mijenjale i evoluirale. Haenlein i sur., (2019) definiraju AI kao sposobnost sustava da točno interpretira vanjske podatke, uči iz tih podataka, te koristi stečena znanja za ostvarivanje specifičnih ciljeva i zadataka kroz fleksibilnu prilagodbu. S druge strane, Poole i Mackworth (2010) umjetnu inteligenciju vide kao polje koje istražuje sintezu i analizu računalnih agenata koji djeluju inteligentno. Ove različite perspektive naglašavaju kako je AI simultano i visoko i nisko rizičan, ovisno o njegovoj primjeni, što dodatno potkrepljuje potrebu za detaljnim razlikovanjem u raspravama o ovoj temi. Objekt (ili osoba) koji djeluje poznat je kao agent. Agent je inteligentan kada:

1. njegovi postupci su prikladni za njegove okolnosti i njegove ciljeve
2. fleksibilan je prema promjeni okruženja i promjeni ciljeva
3. uči iz iskustva i
4. donosi odgovarajuće izbore s obzirom na svoja perceptivna i računalna ograničenja.

Prema definiciji umjetne inteligencije (AI) koju nudi Stuart Russell, umjetna inteligencija se istražuje kao disciplina koja se bavi inteligentnim agentima koji primaju upute iz svoje okoline te djeluju u skladu s njima. U okviru ove definicije razlikuju se različiti pristupi i modeli, uključujući proizvodne sustave, reaktivne agente, logičke planere, neuronske mreže i teorijske sustave odlučivanja, što pokazuje širinu metodologija unutar polja umjetne inteligencije (Russell i Norvig, 1996).

Kroz povijest razvoja umjetne inteligencije (AI) iskristalizirale su se četiri glavne struje razmišljanja koje definiraju različite pristupe ovom polju istraživanja (Russell i Norvig, 1996). Prva se usredotočuje na razvoj računalnih sustava koji bi mogli procesirati misli na sličan način kao ljudi, oponašajući mentalne procese i sposobnosti karakteristične za ljudsku inteligenciju.



Druga stavlja naglasak na simulaciju ljudskog ponašanja kroz akcije koje izvode agenti ili roboti, manje se oslanjajući na procese koji vode do tih akcija. Treća škola razmišljanja usmjerena je na razvoj strojeva koji se ponašaju racionalno, pri čemu su racionalnost i optimalnost ključni elementi koji definiraju njihovo ponašanje. Konačno, četvrta teži stvaranju sustava sposobnih za racionalno razmišljanje, gdje je planiranje i donošenje odluka vođeno maksimalnom efikasnošću u rješavanju problema s kojima se sustav suočava. Ovi različiti pristupi ilustriraju kompleksnost i dubinu polja umjetne inteligencije, kao i širok raspon njenih primjena i tehnologija.

Pojam "umjetna inteligencija" podrazumijeva širok spektar računalnih sustava dizajniranih za razumijevanje okoline, sudjelovanje u kognitivnim procesima te mogućnost učenja kako bi optimalno reagirali u cilju postizanja zadanih ciljeva. Kako navode Gillham i suradnici (2018), ovi sustavi simuliraju aktivnosti koje su tradicionalno zahtijevale ljudsku inteligenciju. Primjena AI varira od usko specijaliziranih zadataka, poput optimizacije potrošnje energije u pametnim mrežama, što se obično kategorizira kao "slaba" ili "uska" umjetna inteligencija, do složenijih sustava kao što su napredni chatbotovi koji imaju sposobnost vođenja složenijih razgovora.

Zajednički element svih oblika AI je njihova sposobnost da se dizajniraju kao inteligentni agenti koji mogu postići ciljeve temeljem percepcije i kognicije. Napredniji oblici AI karakterizira prilagodljivost i sposobnost učenja iz prošlih iskustava, što im omogućuje da osjete, planiraju i djeluju izvan svojih početnih konfiguracija. Takva adaptabilnost čini ih naprednijima od sustava koji se oslanjaju samo na fiksna pravila za određena ponašanja. Međutim, postoje situacije, poput izvođenja specifičnih medicinskih postupaka, gdje je poželjno ograničiti mogućnost učenja AI sustava, budući da su predvidljivost i preciznost od ključne važnosti (Gillham i sur., 2018).

Različiti pristupi unutar polja umjetne inteligencije naglašavaju različite aspekte ovih sistema. Na primjer, razvoj ekspertnih sistema se usredotočuje na stvaranje baza znanja koje ljudi mogu konsultirati za specifične savjete, dok se u razvoju sistema za strojno učenje težište stavlja na stvaranje algoritama koji mogu samostalno otkrivati nova znanja kroz analizu podataka. Ova raznolikost pristupa odražava širinu i dubinu umjetne inteligencije kao disciplinu koja kontinuirano evoluira i prilagođava se različitim aplikacijama i izazovima (Gillham i sur., 2018).

Prema istraživanju koje su proveli Samoili i suradnici (2020), sustavi umjetne inteligencije mogu se definirati isključivo kao softverski, a ponekad i hardverski sistemi, koje je osmislio čovjek. Ovi sustavi su sposobni funkcionirati kako u fizičkim tako i u digitalnim okruženjima s ciljem postizanja zacrtanih ciljeva. Oni se prilagođavaju i razumiju svoje okruženje putem prikupljanja i analize kako strukturiranih tako i nestrukturiranih podataka. Ova sposobnost obrade informacija omogućava im da identificiraju najefikasnije radnje koje trebaju poduzeti kako bi uspješno ostvarili svoje ciljeve.

Umjetna inteligencija koristi razne metode učenja i adaptacije, uključujući primjenu simboličkih pravila ili razvoj numeričkih modela, što joj omogućava da prilagodi svoje ponašanje na temelju ocjene učinka prethodnih akcija na okolinu. Ova sposobnost kontinuirane adaptacije i učenja iz iskustva ključna je za dinamičku prirodu AI sustava i njihovu efikasnost u različitim aplikacijama. Konceptualizacija umjetne inteligencije koju nude Samoili i suradnici (2020) ističu ovu prilagodljivost i evolutivni kapacitet AI sustava, naglašavajući važnost njihove sposobnosti da samostalno optimiziraju svoje postupke kako bi postigli optimalne rezultate u mijenjajućim okolnostima.

## 2.2. Tipovi umjetne inteligencije

Pojam umjetne inteligencije obuhvaća razne pokušaje da se stvore strojevi koji pokazuju određenu razinu inteligencije. Kako se ovo polje razvijalo u posljednja desetljeća, pojavile su se brojne klasifikacije AI sustava temeljena na njihovim različitim karakteristikama i sposobnostima (Russell i Norvig, 1996). Jedna od najčešćih funkcionalnih podjela uključuje tri glavne kategorije: usku AI, općenitu AI i superinteligenciju.

Uska umjetna inteligencija odnosi se na sustave dizajnirane i trenirane za izvršavanje vrlo specifičnih zadataka, poput prepoznavanja govora, rukopisa, preporučivanja proizvoda, chatbotova za korisničku podršku ili odgovaranja na jednostavna pitanja. Iako korisni, ovi sustavi su izrazito ograničeni samo na svoju usku domenu ekspertize. Oni nemaju svijest o sebi niti posjeduju emocionalnu inteligenciju poput ljudi (Jajal, 2021). Današnji virtualni asistenti i alati za obradu prirodnog jezika spadaju u kategoriju "slabe" AI jer su daleko ispod ljudskih

kognitivnih sposobnosti. Oni nisu sposobni za samostalno razmišljanje jer im nedostaje istinska svijest i inteligencija ključne za ljudsku spoznaju.

Primjerice, Siri ne funkcionira sa sviješću, već samo analizira naš govor, unosi ga u tražilicu i vraća rezultate. Zbog toga ne može dati smislene odgovore na apstraktna pitanja poput "koji je smisao života?", već samo vraća internetske poveznice. Međutim, na pitanje poput "koliko je sati?", Siri će dati precizan odgovor jer spada u područje njezine ograničene inteligencije. Za razliku od ljudi, današnja AI jednostavno nema kapacitet za fluidno i fleksibilno razmišljanje kakvo mi posjedujemo. Složeni sustavi poput autonomnih vozila se, ako nisu sastavljeni od više uskih AI sustava, smatraju da imaju lošu AI. S druge strane, opća umjetna inteligencija (AGI) predstavlja hipotetski AI sustav koji bi mogao demonstrirati razine ljudske inteligencije i kognitivnih sposobnosti u gotovo svim područjima (Goertzel 2007). Za razliku od uske AI usmjerene na jednu specifičnu domenu, AGI bi bio sposoban razumjeti kontekst, učiti iz iskustva, planirati, rješavati probleme i prenositi znanje između različitih domena. Iako stvaranje prave AGI još uvijek nije ostvareno, ona ostaje dugoročni cilj mnogih AI istraživača.

Većina trenutnih AI istraživanja usmjerena je na razvoj "uske AI", odnosno programa koji pokazuju inteligenciju u jednoj specijaliziranoj oblasti poput šaha, medicinske dijagnostike, autonomne vožnje, algebarskih izračuna ili dokazivanja matematičkih teorema. Neki od tih specijaliziranih AI programa postigli su zapanjujuće rezultate u svojim industrijama. Međutim, ideja AGI koja se obrađuje u ovom tekstu fundamentalno je drugačija - ona se fokusira na stvaranje softverskog programa sposobnog samostalno rješavati složene probleme u više domena, a koji bi imao vlastite misli, brige, osjećaje, snage, slabosti i preferencije. Opća inteligencija podrazumijeva sposobnost stjecanja širokog spektra znanja i vještina, neovisno o specifičnim domenama. Prema Goertzelu (2007), ta sposobnost učenja je vrlo prilagodljiva, samousmjerena i usmjerena na postizanje određenih ciljeva. Učenje takve autonomne inteligencije mora se odvijati nenamjerno, bez nadzora, kroz dvosmjernu interakciju s okolinom, uključujući istraživanje i eksperimentiranje.

Opća kognitivna sposobnost suprotna je inherentnim specijalizacijama poput govora, a upravo ona omogućuje entitetu stjecanje gotovo beskonačne raznolikosti novih specijaliziranih vještina. Općenito inteligentan sustav može učiti nove stvari, unaprijediti postojeće znanje i vještine te ih učinkovitije primijeniti u odgovarajućem kontekstu od drugih sustava. Pritom je ključno da se znanje stječe i pohranjuje na način prikladan karakteru podataka i ciljevima

(Churchland, 1996). Opći AI sustav, ako posjeduje odgovarajući skup temeljnih sposobnosti, trebao bi moći naučiti prepoznavati i kategorizirati širok raspon novih perceptivnih obrazaca dobivenih kroz različita osjetila i okruženja. Uz povratni mehanizam, takav sustav također bi trebao biti u stanju samostalno naučiti odgovarajuće, ciljano usmjerene odgovore na različite ulazne kontekste.

Konačno, umjetna superinteligencija odnosi se na hipotetski AI sustav koji bi značajno nadmašio ljudske kognitivne sposobnosti u gotovo svakom pogledu - od zaključivanja, planiranja, učenja do apstraktnog mišljenja i kreativnosti. Iako je superinteligencija još uvijek znanstvena fantastika, neki futuristi predviđaju da bi se mogla ostvariti do kraja ovog stoljeća uz kontinuirani eksponencijalni rast u računalnoj snazi i tehnikama strojnog učenja (Bostrom, 2014).

Stvaranje umjetne superinteligencije predstavlja svojevrsan Sveti gral u polju umjetne inteligencije jer bi rezultiralo najnaprednijim oblikom inteligencije ikad viđenim. Takav sustav mogao bi jednog dana oponašati ili čak nadmašiti zamršenu multidisciplinarnu ljudsku inteligenciju. Osim toga, zahvaljujući svojoj ogromnoj memoriji, brzini obrade podataka i sposobnosti donošenja odluka, umjetna bi superinteligencija s vremenom značajno nadmašila ljudske sposobnosti u svim zadacima (Kurzweil, 2017). Razvoj ovakve superinteligencije čini "tehnološku singularnost" realnom mogućnošću u budućnosti.

No to otvara pitanje - koja bi onda bila uloga nas ljudi ako strojevi postanu napredniji u svakom aspektu i nastave se eksponencijalno razvijati? Štoviše, ako ti strojevi ne samo da su superiorniji, već imaju i svijest, što bi ih spriječilo da svoj razvoj i nastavak postojanja stave ispred našeg? S druge strane, futurist Ray Kurzweil (2017) smatra da bismo mogli koegzistirati s umjetnom superinteligencijom u svijetu u kojem će njezine sposobnosti samo povećavati naše vlastite.

Predviđa se da će potreban hardver za pokretanje superinteligentnih sustava biti dostupan u sljedećih 20-30 godina. Međutim, razvoj potrebnog softvera predstavljat će veći izazov. Potraga za umjetnom superinteligencijom trenutno je uglavnom teorijska jer nema konkretnih dokaza koji bi poduprli ili opovrgli mogućnost da to postane stvarnost. Web stranica AI Impacts sakupila je uvide i stajališta raznih profesionalaca iz industrije AI-a. Prema njihovim anketama među stručnjacima, srednja procjena pokazuje 10% šansu za postizanje ljudske razine AI do

2020-ih i 50% šansu između 2035. i 2050. (AI Impacts, 2021). Štoviše, mnogi smatraju da će AI na ljudskoj razini biti ostvarena do 2085., uključujući i brojne stručnjake iz znanosti i tehnologije (AI Impacts, 2021).

Osim ove tri osnovne kategorije (uska, opća i superinteligencija), postoje i brojni drugi načini klasifikacije AI sustava. Na primjer, postoji razlika između "jake" i "slabe" AI. Slaba AI odnosi se na usko specijalizirane sustave koji samo simuliraju neke aspekte inteligencije, dok jaka AI opisuje hipotetske buduće strojeve s pravim kognitivnim sposobnostima na razini ljudi. Druga korisna podjela je na "simboličku" i "vezujuću" AI - prva se oslanja na pravila i logiku za reprezentaciju znanja, a druga na statističko učenje iz velikih skupova podataka.

Još jedna važna distinkcija je između "uske" i "široke" AI. Uska AI rješava vrlo određene probleme u uskom području, dok široka AI pokušava riješiti širok raspon problema u raznim domenama, što zahtijeva mnogo veću količinu znanja i kognitivnu fleksibilnost. Većina današnjih komercijalnih AI aplikacija spada u kategoriju uske AI. Nadalje, AI sustavi mogu se klasificirati prema vrsti pristupa koji koriste - neki se oslanjaju na tradicionalne simboličke tehnike poput pretraživanja stabala, logike i pravila, dok drugi koriste pristupe temeljene na populaciji poput genetskih algoritama. Strojno učenje i duboko učenje trenutno dominiraju komercijalnim primjenama AI-a.

S porastom broja različitih AI tehnika i aplikacija, pojavila se potreba za preciznijim taksonomijama koje adekvatno klasificiraju ove sustave. AAAI je nedavno objavio preliminarni izvještaj koji predlaže novu AI taksonomiju s 5 glavnih grana: percepcija i interpretacija podataka, razumijevanje jezika, računarsko zaključivanje, djelovanje u fizičkom svijetu i integracija (AAAI, 2022). Ova taksonomija nastoji klasificirati AI sustave na temelju njihovih sposobnosti i funkcionalnosti umjesto na temelju korištenih pristupa i tehnika.

Bez obzira na točnu taksonomiju, jasno je da umjetna inteligencija predstavlja izuzetno široko i raznoliko područje u stalnom razvoju. Dok se mnoge rasprave fokusiraju na pitanje hoće li i kada će opća ili superinteligencija biti ostvarena, većina komercijalnih primjena i dalje spada u kategoriju uske AI. Međutim, kontinuirani napredak u područjima poput strojnog učenja, obrade prirodnog jezika i računarstva u oblaku omogućuje sve širi spektar inteligentnih aplikacija. Razvoj robustnih okvira za testiranje, verifikaciju i valjanost AI sustava bit će ključan za osiguravanje njihove sigurnosti i korisnosti.

Iz perspektive zapošljavanja, rast i širenje AI aplikacija vjerojatno će transformirati mnoga zanimanja i industrije. Iako će neka radna mjesta nestati, AI će stvoriti potrebu za novim ulogama poput inženjera podataka, stručnjaka za strojno učenje i developera robotike. Razvoj etičkih smjernica i javnih politika bit će presudan za osiguravanje da dobrobit društva bude u središtu ovih tehnoloških promjena. Klasifikacija AI sustava složeno je i multidimenzionalno pitanje. Iako postoje brojni načini kategorizacije temeljeni na različitim značajkama, funkcionalna podjela na usku, opću i superinteligenciju i dalje pruža koristan pregled područja. Dok većina trenutnih komercijalnih primjena potpada pod usku AI, ambiciozniji oblici opće i super inteligencije ostaju tehnološka fantazija barem još neko vrijeme. Međutim, ovo brzo evoluirajuće polje zahtijeva stalnu pažnju kako bi se osiguralo da razvoj i primjena ovih tehnologija bude etička, sigurna i korisna za društvo u cjelini.

### 2.3. Utjecaj umjetne inteligencije na poslovne procese

Korištenje umjetne inteligencije (AI) u poslovnom sektoru sve više postaje popularno, s ciljem povećanja učinkovitosti i profitabilnosti poslovnih procesa. Primjećuje se da kompanije različitih veličina, od velikih multinacionalnih korporacija do malih i srednjih poduzeća, sve više prihvaćaju umjetnu inteligenciju kao alat za transformaciju svojih konvencionalnih operacija. Umjetna inteligencija igra ključnu ulogu u unapređenju automatizacije poslovnih procesa, što podrazumijeva korištenje tehnoloških rješenja za automatiziranje rutinskih i ponavljajućih zadataka. Primjerice, dokumentirano je kako su organizacije poput NASA-e implementirale sustave robotske automatizacije procesa (RPA) u različitim segmentima, uključujući upravljanje ljudskim resursima i financije. Prema istraživanju Davenporta i Ronankija (2018), ovi sustavi značajno umanjuju potrebu za ljudskom intervencijom zahvaljujući svojoj jednostavnosti i efikasnosti.

Široka primjena algoritama dubokog i strojnog učenja u svrhu izvlačenja korisnih informacija iz velikih skupova podataka također je zapažena. Na primjer, Amazon koristi sustav za preporuke koji se temelji na podacima kako bi optimizirao ponude za svoje kupce, što doprinosi boljem razumijevanju preferencija korisnika. Davenport i Ronanki (2018) ističu da kompanije koje uspijevaju dublje razumjeti preferencije svojih kupaca bilježe više razine zadovoljstva korisnika. Nadalje, upotreba tehnologija za obradu prirodnog jezika u realnom vremenu

predstavlja značajan napredak u primjeni umjetne inteligencije, omogućavajući efikasniju komunikaciju i analizu podataka. Emocionalna inteligencija integrirana s algoritmima strojnog učenja omogućava sisteme poput chatbotova u korisničkim službama da pružaju interakcije visoke kvalitete, do te mjere da korisnici često nisu ni svjesni da komuniciraju s računalnim programima, umjesto s ljudima.

Tvrtke sve više koriste umjetnu inteligenciju za unaprjeđenje ključnih aspekata svog poslovanja, uključujući upravljanje zalihama, dizajniranje izgleda trgovina i predviđanje potrošačkih preferencija. Ovakva primjena AI-a doprinosi optimizaciji prodajnih i marketinških odjela, omogućavajući bolje ciljanje i razumijevanje potrošačkih motivacija. Osim toga, AI tehnologije, poput Alibabinog AI-CopyWriter-a, pokazuju sposobnost stvaranja kvalitetnih reklamnih materijala, što je značajan napredak u području kreiranja sadržaja. Integracija umjetne inteligencije s tradicionalnom robotikom rezultirala je poboljšanom suradnjom robota i ljudskih radnika, što je dovelo do sigurnijih i efikasnijih operacija (Agrawal i sur., 2019).

U kontekstu zapošljavanja, tvrtke se suočavaju s velikim brojem prijavi za posao koristeći AI rješenja, kao što su ona koja nudi Pymetrics, što ubrzava proces zapošljavanja i smanjuje subjektivnost u odabiru kandidata. Također, sigurnosni sustavi vođeni umjetnom inteligencijom pokazali su se superiornijima u odnosu na tradicionalne metode u detekciji prijevara, zahvaljujući svojoj sposobnosti učenja i prilagodbe (Agrawal i sur., 2019).

Postoji predviđanje da će tvrtke uskoro početi uvoditi specijalizirane čipove za obradu algoritama u stvarnom vremenu, što bi omogućilo široku upotrebu umjetne inteligencije u lancima opskrbe. Osim toga, kako se tehnologija 5G nastavlja širiti, umjetna inteligencija bi mogla postati ključna komponenta u funkcioniranju Interneta stvari (IoT). Nadalje, kombinacija umjetne inteligencije i decentraliziranih blockchain tehnologija privlači značajan interes, sugerirajući potencijal za revolucionarne inovacije u tom području. Ova široka primjena i budući izgledi umjetne inteligencije u poslovnim procesima nagovještavaju duboke transformacije u načinu kako tvrtke djeluju globalno.

## 2.4. Primjeri upotrebe umjetne inteligencije

Umjetna inteligencija (AI) ostvarila je značajne promjene u mnogim industrijskim sektorima, uključujući zdravstvo, e-trgovinu, robotiku, financije, prepoznavanje lica, marketing i društvene medije, gdje omogućava automatizaciju složenih zadataka, poboljšava procese donošenja odluka i transformira korisnička iskustva. U zdravstvu, AI olakšava administrativne i kliničke procese, što rezultira boljom organizacijom podataka pacijenata i povećanjem efikasnosti liječnika. Napredna obrada prirodnog jezika (NLP) koristi se za upravljanje medicinskim zapisima, dok algoritmi za računalni vid i konvolucijske neuronske mreže (CNN) omogućuju preciznu interpretaciju složenih medicinskih slika, poput MRI skenova. Također, umjetna inteligencija povećava preciznost, sigurnost i brzinu robotskih operacija (Lanfranco i sur., 2004).

U industriji e-trgovine, AI nudi konkurentsku prednost personalizirajući kupovno iskustvo i pružajući izvrsnu korisničku službu. Sistemi za preporuke vođeni AI analiziraju ponašanje korisnika i njihove preference kako bi ponudili personalizirane proizvode, čime povećavaju angažman i lojalnost kupaca. Virtualni pomoćnici i chatbotovi vođeni umjetnom inteligencijom pružaju gotovo ljudsku interakciju, poboljšavajući online iskustvo kupovine. Osim toga, umjetna inteligencija pomaže u otkrivanju i sprječavanju prijevara, analizirajući obrasce korištenja kako bi se identificirale lažne recenzije i prijekave s kreditnim karticama.

Robotika, koja je neophodna od svog izuma, sada postaje još sposobnija i efikasnija zahvaljujući integraciji s umjetnom inteligencijom. Roboti opremljeni sposobnostima računalnog vida mogu bolje navigirati i reagirati na svoje okruženje, omogućavajući napredniju i autonomniju operativnu funkcionalnost (LeCun i sur., 2015). Ove inovacije dokazuju kako umjetna inteligencija može biti temeljna za unaprjeđenje širokog spektra industrijskih aplikacija, čime se osigurava njezina sve veća važnost i prisutnost u različitim sektorima.

Umjetna inteligencija (AI) revolucionirala je i financijski sektor, pružajući značajne inovacije u procesima donošenja kreditnih odluka, kvantitativnom trgovanju i upravljanju financijskim rizicima. Prema Ameenu i suradnicima (2022), AI je značajno pojednostavio operacije unutar financijske industrije kroz poboljšanja u procjeni rizika, detekciji i upravljanju prijevarama, financijskom savjetovanju i automatizaciji trgovačkih procesa.

Što se tiče marketinga, AI tehnologije imaju široku primjenu, omogućavajući marketing sustavima da analiziraju podatke i identificiraju ekonomske trendove i obrasce ponašanja



publike koji mogu imati utjecaj na marketinške strategije. AI poboljšava ciljanje oglasa kroz sofisticiranu analizu ponašanja korisnika i prepoznavanje uzoraka, što omogućava dostavu prilagođenih reklamnih sadržaja. Marketing stručnjaci koriste AI za brzo donošenje odluka temeljeno na taktičkim podacima prikupljenim u stvarnom vremenu, što značajno povećava efikasnost i učinkovitost marketinških kampanja.

Na polju društvenih medija, AI analizira velike količine podataka kako bi razumio ponašanje korisnika, prateći nestrukturirane komentare i personalizirajući korisnička iskustva te identificirajući potencijalne krize. Također, AI koristi demografske podatke i korisničke aktivnosti za kreiranje angažiranog sadržaja. Primjerice, umjetna inteligencija upravlja Instagramovom karticom Explore, prilagođavajući sadržaj na temelju interesa korisnika i profila koje prate, čime se osigurava relevancija i privlačnost sadržaja predstavljenog svakom pojedincu.

Istraživanje autonomnih vozila sve više koristi umjetnu inteligenciju (AI) u automobilskoj industriji. AI omogućava analizu podataka s različitih senzora za identifikaciju okoline, predviđanje promjena u prometu, donošenje odluka i precizno upravljanje vozilom. Također, sustavi za upravljanje prometom koriste AI za optimizaciju prometnih signala i ruta, čime se smanjuje prometna gužva i povećava efikasnost prijevoza.

U poljoprivrednom sektoru, AI igra ključnu ulogu u automatizaciji i poboljšanju učinkovitosti. Strojni algoritmi koriste se za analizu poljoprivrednih podataka radi predviđanja uspjeha određenih usjeva, što doprinosi komercijalnoj uspješnosti. Osim toga, AI omogućava praćenje i kontrolu upotrebe resursa poput vode i gnojiva, čime se poljoprivreda čini održivijom (Kamilaris i sur., 2017).

Tehnologije AI također proširuju granice svemirskih istraživanja. NASA-in rover Perseverance i druga autonomna vozila koriste AI za navigaciju i istraživanje Marsa, donošenje odluka na terenu i slanje podataka nazad na Zemlju. AI također analizira velike skupove teleskopskih podataka u potrazi za novim uzorcima i fenomenima koji bi inače mogli promaknuti znanstvenicima. U glazbenoj industriji, AI se koristi za stvaranje novih pjesama, analizu korisničkih preferencija te preporuku i prilagodbu glazbenog sadržaja korisnicima na platformama poput Spotifyja, SoundClouda i Pandya. Ove tehnologije omogućuju bolje korisničko iskustvo kroz personalizirane preporuke.

AI također igra važnu ulogu u obrazovanju, prilagođavajući se individualnim potrebama učenika, identificirajući područja u kojima učenici imaju poteškoće i omogućavajući učiteljima da automatiziraju administrativne zadatke poput ocjenjivanja i planiranja (Surden, 2014).

Pravne usluge također koriste AI za analizu velikih pravnih dokumenata, podršku u pravnim istraživanjima i predviđanje sudske prakse, s aplikacijama poput ROSS Intelligence i LegalRobot koji nude pravnu pomoć zasnovanu na AI. U industriji videoigara, AI unosi novu dinamiku pružajući likovima inteligentno i nepredvidivo ponašanje, čime se igre čine realističnijima i interaktivnijima (Yannakakis i Togelius, 2018). Općenito, umjetna inteligencija transformira brojne industrije, od poljoprivrede do svemirskih istraživanja, obrazovanja, pravnih usluga i logistike, pružajući temelj za inovacije i poticanje rasta u tim sektorima.

## 2.5. Budući trendovi i izazovi u razvoju umjetne inteligencije

Umjetna inteligencija (AI) ključni je pokretač Četvrte industrijske revolucije. Njezin utjecaj vidljiv je u domovima, poslovnim i javnim prostorima. U obliku robota, uskoro će upravljati automobilima, opskrbljivati skladišta i brinuti se za mlade i starije. Obećava rješenje nekih od najvažnijih društvenih problema, ali predstavlja i izazove poput nejasnih algoritama "crne kutije", neetične upotrebe podataka i moguće zamjene za poslove. Kako brzi napredak u strojnom učenju povećava opseg i razmjor primjene AI-a u svim aspektima svakodnevnog života, a tehnologija može učiti i mijenjati se sama, suradnja više dionika postaje nužna za optimizaciju odgovornosti, transparentnosti, privatnosti i nepristranosti kako bi se stvorilo povjerenje.

Prilike uključuju poboljšanje ekonomskih ishoda i produktivnosti: poput drugih tehnoloških napredaka u prošlosti, umjetna inteligencija će poboljšati brzinu i učinkovitost proizvodnje. Međutim, istaknuto je da će mjerenje utjecaja AI-a biti teško i trenutno ne postoje mehanizmi koji bi točno mjerili njezin utjecaj. Poboljšanje ili asistiranje u donošenju ljudskih odluka također se navodi kao prilika, gdje AI omogućava svojim korisnicima integraciju i otkrivanje trendova ili anomalija skrivenih unutar ogromnih i raznolikih skupova podataka. Donositelji politika mogu koristiti sustave AI-a za stvaranje politika temeljenih na podacima, iako

validacija i potencijalna pristranost takvih sustava još nisu dobro razumjeli. Također se navodi poboljšanje rješavanja problema: trenutni napredak u istraživanju AI-a obećava sve veću primjenu tehnologije na društvene izazove, istovremeno minimizirajući teret regulatornog nadzora za vladu i regulirane subjekte (Ariffin i sur., 2023).

Izazovi uključuju prepreke pri prikupljanju i dijeljenju podataka: sustavi umjetne inteligencije koji koriste različite izvore podataka mogu naići na poteškoće u pristupanju i integraciji podataka iz izvora koji se razlikuju u regulatornoj dostupnosti, potpunosti i općoj kvaliteti tih podataka. Ograničen pristup računalnim resursima i ljudskom kapitalu također predstavlja izazov, gdje razvijatelji, istraživači i implementatori u različitim vladinim organizacijama ili agencijama mogu imati poteškoće u dobivanju i financiranju računalne snage i intenzivnih potreba za talentima koje zahtijevaju sustavi AI-a (Dagnaw, 2020).

Pravne i regulatorne prepreke su istaknute kao izazov, gdje brzi napredak i primjena sustava AI-a na neke načine premašuju regulatorni okvir za upravljanje kako i gdje se ti sustavi trebaju učinkovito i sigurno koristiti u brojnim primjenama. Potrebna je nova tehnološka stručnost unutar vlade kako bi se osiguralo da je politika za AI ažurirana i primjerena za tehnologiju. Razvijanje etičkih, objašnjivih i prihvatljivih aplikacija AI-a također predstavlja izazov, jer kako sustavi umjetne inteligencije poboljšavaju i sve više nadmašuju ljudske sposobnosti, važno je da se radnje i odluke proizašle iz tih sustava mogu smatrati odgovornima kao i ljudski donositelji odluka koje pomažu i/ili zamjenjuju (Dagnaw, 2020).

Brza tehnološka promjena, robotika, veliki podaci, umjetna inteligencija i povezani objekti nude nove mogućnosti proizvođačima. No, donose i izazove industriji proizvodnje. Na primjer, u Quebecu je industrija proizvodnje u riziku, što predstavlja gotovo 800.000 izravnih i neizravnih radnih mjesta. Kako bi uspješno prešla na Industriju 4.0, proizvodna tvrtka mora istražiti nove potrebne vještine i potrebu za kvalificiranim osobljem. Najtraženije industrijske vještine u 4.0 uključuju upravljanje podacima, sigurnost podataka, interakciju čovjeka i stroja, dizajn korisničkog sučelja, razvoj softvera, programiranje, znanost o podacima i analitiku. Glavni izazov s kojim se tvrtka suočava je obučavanje zaposlenika i zapošljavanje novih resursa. Radi se o pronalaženju najprikladnijeg pristupa za tvrtku kako bi uspješno preoblikovala vrijednosni lanac i održala ili izgradila svoje konkurentske prednosti. Studije provedene u Njemačkoj i Sjedinjenim Državama pokazale su da većina industrijskih radnika nema vještine potrebne za Industriju 4.0.

Sigurnost podataka predstavlja zabrinutost za sve tvrtke koje su odlučile prijeći na Industriju 4.0. Umnožavanje podataka i sustava unutar tvrtke ističe važnost aspekta računalne sigurnosti. Kada su tehnologije bile povezane s unutarnjom mrežom i centralizirane u istoj zgradi, bilo je lakše osigurati sve. Dolazak mnoštva povezanih objekata, često premještenih i dostupnih putem interneta, sada nameće upravljanje kibernetičkom sigurnošću. Stoga je ključno integrirati elemente kibernetičke sigurnosti u implementaciju IT infrastrukture tvrtke. Nove tehnologije su stalno povezane s ulaganjima. Industrija 4.0 posebno označava temeljnu promjenu za tvrtke koja zahtijeva znatna financijska sredstva. Nažalost, većina tvrtki nema takav veliki iznos sredstava i ne uspijeva se implementirati u Industriju 4.0.

### 3. REGULATORNI OKVIR ZA UPORABU UMJETNE INTELIGENCIJE U POSLOVANJU - AKT O UMJETNOJ INTELIGENCIJI

#### 3.1. Pravne osnove, načelo supsidijarnosti i proporcionalnosti

Pravna osnova za prijedlog prije svega je Članak 114 Ugovora o funkcioniranju Europske unije (TFEU), koji omogućuje donošenje mjera za osiguranje uspostave i funkcioniranja unutarnjeg tržišta.

Prijedlog se smatra ključnim dijelom strategije EU za jedinstveno digitalno tržište. Primarni cilj ovog prijedloga je osigurati pravilno funkcioniranje unutarnjeg tržišta uspostavom usklađenih pravila posebno u vezi s razvojem, stavljanjem na tržište Unije i korištenjem proizvoda i usluga koji koriste tehnologije umjetne inteligencije ili se pružaju kao samostalni sustavi umjetne inteligencije. Razmatranje nacionalnih pravila za osiguranje sigurnosti umjetne inteligencije i njezin razvoj i korištenje u skladu s obvezama temeljnih prava već se razmatra u nekim državama članicama. To će vjerojatno dovesti do dva glavna problema:

- i) fragmentacije unutarnjeg tržišta po bitnim elementima što se tiče posebno zahtjeva za proizvode i usluge umjetne inteligencije, njihovo korištenje, odgovornost i nadzor od strane javnih vlasti,
- ii) znatnog smanjenja pravne sigurnosti za pružatelje i korisnike sustava umjetne inteligencije u vezi s tim kako će se postojeća i nova pravila primjenjivati na te sustave u Uniji. S obzirom na široku distribuciju proizvoda i usluga preko granica, ova dva problema mogu se najbolje riješiti kroz zakonodavstvo EU koje usklađuje (Europska komisija, 2021).

Doista, prijedlog definira zajedničke obvezne zahtjeve primjenjive na dizajn i razvoj određenih sustava umjetne inteligencije prije nego što se stave na tržište, koji će se dodatno operacionalizirati kroz usklađene tehničke standarde. Prijedlog se također bavi situacijom nakon što sustavi umjetne inteligencije budu stavljeni na tržište usklađivanjem načina na koji se provode ex-post kontrole.

Osim toga, s obzirom na to da prijedlog sadrži određena specifična pravila o zaštiti pojedinaca u vezi s obradom osobnih podataka, posebno ograničenja upotrebe sustava umjetne inteligencije za 'stvarno-vremensko' daljinsko biometrijsko prepoznavanje u javno dostupnim prostorima u svrhu provođenja zakona, prikladno je temeljiti ovu regulativu, što se tiče tih specifičnih pravila, na Članku 16 TFEU.

Priroda umjetne inteligencije, koja često ovisi o velikim i raznolikim skupovima podataka te može biti ugrađena u bilo koji proizvod ili uslugu koji slobodno kruže unutar unutarnjeg tržišta, zahtijeva da ciljevi ovog prijedloga ne mogu biti učinkovito postignuti samostalno od strane država članica. Štoviše, nastajanje mozaika potencijalno različitih nacionalnih pravila ometat će besprijekorno kruženje proizvoda i usluga vezanih uz sustave umjetne inteligencije diljem EU te će biti neučinkovito u osiguravanju sigurnosti i zaštite temeljnih prava i vrijednosti Unije među različitim državama članicama. Nacionalni pristupi u rješavanju problema stvorit će samo dodatnu pravnu nesigurnost i prepreke te će usporiti tržišno prihvaćanje umjetne inteligencije (Europska komisija, 2021).

Ciljevi ovog prijedloga bolje se mogu postići na razini Unije kako bi se izbjegla daljnja fragmentacija Jedinog tržišta u potencijalno kontradiktorne nacionalne okvire koji sprječavaju slobodno kruženje roba i usluga koje ugrađuju umjetnu inteligenciju. Čvrst europski regulatorni okvir za pouzdanu umjetnu inteligenciju također će osigurati ravnopravne uvjete i zaštititi sve ljude, dok će ojačati konkurentnost Europe i industrijsku osnovu u području umjetne inteligencije. Samo zajedničko djelovanje na razini Unije također može zaštititi digitalni suverenitet Unije i iskoristiti njezine alate i regulatorne ovlasti za oblikovanje globalnih pravila i standarda.

Prijedlog se temelji na postojećim pravnim okvirima te je proporcionalan i nužan za postizanje svojih ciljeva, budući da se primjenjuje pristup zasnovan na riziku i nameće regulatorna opterećenja samo kada je vjerojatno da će sustav umjetne inteligencije predstavljati visoke

rizike za temeljna prava i sigurnost. Za druge, sustave umjetne inteligencije koji nisu visokog rizika, nameću se samo vrlo ograničene obveze transparentnosti, na primjer u smislu pružanja informacija za označavanje korištenja sustava umjetne inteligencije pri interakciji s ljudima. Za sustave umjetne inteligencije visokog rizika, zahtjevi visokokvalitetnih podataka, dokumentacije i praćenja, transparentnosti, nadzora ljudi, točnosti i robustnosti, strogo su nužni za ublažavanje rizika za temeljna prava i sigurnost koje predstavlja umjetna inteligencija, a koji nisu pokriveni drugim postojećim pravnim okvirima. Usklađeni standardi i podrška u obliku smjernica i alata za usklađenost pomoći će pružateljima i korisnicima da se usklade s zahtjevima navedenim u prijedlogu i minimiziraju svoje troškove. Troškovi koje snose operatori proporcionalni su ciljevima koji se postižu te ekonomskim i reputacijskim koristima koje operatori mogu očekivati od ovog prijedloga (Europska komisija, 2021).

Izbor uredbe kao pravnog instrumenta opravdan je potrebom za jedinstvenom primjenom novih pravila, kao što su definicija umjetne inteligencije, zabrana određenih štetnih praksi omogućenih umjetnom inteligencijom te klasifikacija određenih sustava umjetne inteligencije. Izravna primjenjivost uredbe, u skladu s Člankom 288 TFEU, smanjit će pravnu fragmentaciju i olakšati razvoj jedinstvenog tržišta za zakonite, sigurne i pouzdane sustave umjetne inteligencije. To će postići posebno uvođenjem usklađenog skupa temeljnih zahtjeva u pogledu sustava umjetne inteligencije klasificiranih kao visokorizični te obvezama za pružatelje i korisnike tih sustava, poboljšavajući zaštitu temeljnih prava i pružajući pravnu sigurnost kako za operatere tako i za potrošače. Istovremeno, odredbe uredbe nisu pretjerano propisivačke i ostavljaju prostor za različite razine djelovanja država članica za elemente koji ne podrivaju ciljeve inicijative, posebno unutarnju organizaciju sustava nadzora tržišta i prihvaćanje mjera za poticanje inovacija.

### 3.2. Zabranjeni AI sustavi i visoko rizični AI sustavi

U okviru novih regulacija koje se odnose na umjetnu inteligenciju, određene prakse su zabranjene kako bi se zaštitili pojedinci i grupe od potencijalne zloupotrebe. Zabranjuje se stavljanje na tržište, puštanje u rad ili korištenje sustava umjetne inteligencije koji koriste subliminalne tehnike koje prelaze granice svijesti pojedinca ili namjerno manipulativne i obmanjujuće tehnike. Cilj ili učinak ovih tehnika je materijalno iskrivljavanje ponašanja pojedinca ili skupine osoba na način da znatno otežava donošenje informiranih odluka, čime se

osobe navode na odluke koje inače ne bi donijele, a koje mogu prouzročiti značajnu štetu (Europska komisija, 2024).

Također se zabranjuje korištenje sustava umjetne inteligencije koji iskorištavaju ranjivosti pojedinaca ili specifičnih skupina zbog njihove dobi, invaliditeta ili posebne društvene ili ekonomske situacije. Ovi sustavi mogu dovesti do materijalnog iskrivljenja ponašanja na način koji može prouzročiti značajnu štetu.

Posebno je zabranjeno stavljanje na tržište ili puštanje u rad biometrijskih sustava kategorizacije koji kategoriziraju osobe na temelju njihovih biometrijskih podataka kako bi se zaključilo ili pretpostavilo njihovo rasno ili etničko porijeklo, političke stavove, članstvo u sindikatu, vjerska ili filozofska uvjerenja, seksualni život ili seksualnu orijentaciju. Ova zabrana ne obuhvaća označavanje ili filtriranje zakonito stečenih biometrijskih skupova podataka, poput slika, na temelju biometrijskih podataka ili kategorizaciju biometrijskih podataka u području provođenja zakona (Europska komisija, 2024).

Dodatno, zabranjuje se korištenje sustava umjetne inteligencije za ocjenjivanje ili klasifikaciju prirodnih osoba ili skupina na temelju njihovog društvenog ponašanja ili poznatih, pretpostavljenih ili predviđenih osobnih ili osobnostnih karakteristika tijekom određenog vremenskog razdoblja. Društveno bodovanje koje proizlazi iz takve ocjene može dovesti do štetnog ili nepovoljnog tretmana pojedinaca ili cijelih skupina u društvenim kontekstima koji nisu povezani s kontekstima u kojima su podaci prvotno generirani ili prikupljeni, ili do štetnog ili nepovoljnog tretmana koji nije opravdan ili je nesrazmjeran u odnosu na njihovo društveno ponašanje ili njegovu težinu.

Zabranjuje se uporaba sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona, osim ako i u mjeri u kojoj je takva upotreba strogo nužna za jedan od sljedećih ciljeva:

- ciljana potraga za specifičnim žrtvama otmica, trgovine ljudima i seksualnog iskorištavanja ljudi, kao i potraga za nestalim osobama;
- sprječavanje specifične, značajne i neposredne prijetnje životu ili fizičkoj sigurnosti prirodnih osoba ili stvarne i trenutne ili stvarne i predvidive prijetnje terorističkog napada;



- lokalizacija ili identifikacija osobe za koju se sumnja da je počinila kazneno djelo, u svrhu provođenja kaznenog istraživanja, kaznenog progona ili izvršenja kaznene sankcije za djela navedena u Prilogu IIa i kažnjiva u dotičnoj državi članici kaznom zatvora ili mjerom pritvora na najmanje četiri godine. Ovaj stavak ne utječe na odredbe članka 9. GDPR-a o obradi biometrijskih podataka u svrhe koje nisu provođenje zakona (Europska komisija, 2024).

Također se zabranjuje stavljanje na tržište, puštanje u rad za ovu specifičnu svrhu ili upotreba AI sustava za procjenu rizika prirodne osobe kako bi se procijenio ili predvidio rizik da prirodna osoba počinu kazneno djelo, temeljeno isključivo na profiliranju prirodne osobe ili procjeni njihovih osobnih karakteristika i osobina. Ova zabrana ne odnosi se na AI sustave koji se koriste za podršku ljudskoj procjeni uključenosti osobe u kriminalnu aktivnost, koja je već temeljena na objektivnim i provjerljivim činjenicama izravno povezanim s kriminalnom aktivnošću;

Zabranjuje se stavljanje na tržište, puštanje u rad za ovu specifičnu svrhu ili upotreba AI sustava koji stvaraju ili proširuju baze podataka za prepoznavanje lica putem neselektivnog skidanja slika lica s interneta ili snimaka s nadzornih kamera. Zabranjuje se stavljanje na tržište, puštanje u rad za ovu specifičnu svrhu ili upotreba AI sustava za zaključivanje o emocijama prirodne osobe u područjima radnog mjesta i obrazovnih ustanova, osim u slučajevima kada je upotreba AI sustava namijenjena medicinskim ili sigurnosnim razlozima.

Članak 1a ne utječe na zabrane koje se primjenjuju kada praksa umjetne inteligencije krši drugi zakon Unije. Upotreba sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona za bilo koji od ciljeva navedenih u stavku 1 točki (d) može se primijeniti samo za potvrđivanje identiteta određene ciljane osobe. Pri tome se moraju uzeti u obzir sljedeći elementi:

- priroda situacije koja je dovela do moguće upotrebe, posebno ozbiljnost, vjerojatnost i opseg štete koja bi nastala u odsustvu upotrebe sustava;
- posljedice upotrebe sustava za prava i slobode svih zainteresiranih osoba, posebno ozbiljnost, vjerojatnost i opseg tih posljedica (Europska komisija, 2024).

Osim toga, upotreba sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona mora biti u skladu s potrebnim i

proporcionalnim zaštitnim mjerama i uvjetima u skladu s nacionalnim zakonodavstvima koja odobravaju njegovu upotrebu, posebno s obzirom na vremenska, geografska i osobna ograničenja. Upotreba sustava može biti odobrena samo ako je nadležna tijela za provođenje zakona izvršila procjenu utjecaja na temeljna prava kako je predviđeno člankom 29a i registrirala sustav u bazi podataka prema članku 51. Međutim, u slučajevima opravdane hitnosti, upotreba sustava može započeti bez registracije, pod uvjetom da se registracija dovrši bez nepotrebnog odgađanja.

Prema stavcima 1, točka (d) i 2, svaka upotreba sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona mora biti predmet prethodne autorizacije koju dodjeljuje sudsko tijelo ili neovisno administrativno tijelo čija je odluka obvezujuća u državi članici u kojoj će se upotreba odvijati, izdana na temelju obrazloženog zahtjeva i u skladu s detaljnim pravilima nacionalnog zakona navedenima u stavku 4. Međutim, u opravdanoj situaciji hitnosti, upotreba sustava može započeti bez autorizacije, pod uvjetom da se takva autorizacija zatraži bez nepotrebnog odgađanja, najkasnije unutar 24 sata. Ako se takva autorizacija odbije, njegova upotreba mora se odmah zaustaviti i svi podaci, kao i rezultati i izlazi te upotrebe, moraju se odmah odbaciti i izbrisati (Europska komisija, 2024).

Nadležno sudsko tijelo ili neovisno administrativno tijelo čija je odluka obvezujuća može odobriti autorizaciju samo ako je na temelju objektivnih dokaza ili jasnih pokazatelja uvjeren da je upotreba spornog sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju nužna i proporcionalna za postizanje jednog od ciljeva navedenih u stavku 1, točka (d), kako je navedeno u zahtjevu, i posebno ostaje ograničena samo na ono što je strogo nužno u pogledu vremenskog razdoblja, geografskog i osobnog opsega. Pri odlučivanju o zahtjevu, nadležno sudsko tijelo ili neovisno administrativno tijelo čija je odluka obvezujuća mora uzeti u obzir elemente navedene u stavku 2. Mora se osigurati da nijedna odluka koja proizvodi nepovoljan pravni učinak na osobu ne može biti donesena isključivo na temelju izlaza iz daljinskog biometrijskog identifikacijskog sustava (Europska komisija, 2024).

Svaka upotreba sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona mora biti prijavljena relevantnom tijelu za nadzor tržišta i nacionalnom tijelu za zaštitu podataka u skladu s nacionalnim pravilima

navedenim u stavku 4. Obavijest mora sadržavati minimalno informacije navedene pod stavkom 5 i ne smije uključivati osjetljive operativne podatke.

Država članica može odlučiti omogućiti potpuno ili djelomično odobrenje korištenja sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima u svrhu provođenja zakona, unutar granica i pod uvjetima navedenim u stavcima 1, točka (d), 2 i 3. Zainteresirane države članice trebaju u svojim nacionalnim zakonima utvrditi potrebna detaljna pravila za zahtjev, izdavanje i korištenje, kao i nadzor i izvještavanje u vezi s autorizacijama navedenim u stavku 3. Ta pravila također moraju precizirati u odnosu na koje ciljeve navedene u stavku 1, točka (d), uključujući koje kaznene djela navedene u točki (iii) nadležna tijela mogu biti ovlaštena koristiti te sustave u svrhu provođenja zakona.

Države članice moraju obavijestiti Komisiju o tim pravilima najkasnije 30 dana nakon njihova usvajanja. Države članice mogu uvesti, u skladu s pravom Unije, strožija zakonodavstva o korištenju daljinskih biometrijskih sustava identifikacije. Nacionalna tijela za nadzor tržišta i nacionalna tijela za zaštitu podataka država članica koje su obaviještene o korištenju sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima za svrhe provođenja zakona u skladu sa stavkom 3a, moraju Komisiji podnijeti godišnja izvješća o takvoj upotrebi. U tu svrhu, Komisija će pružiti državama članicama i nacionalnim tijelima za nadzor tržišta i zaštitu podataka obrazac, uključujući informacije o broju odluka koje donose nadležna sudbena tijela ili neovisna administrativna tijela čija je odluka obvezujuća na temelju zahtjeva za autorizacijama u skladu sa stavkom 3 i njihovom rezultatu.

Komisija će objavljivati godišnja izvješća o korištenju sustava za 'stvarno-vremensku' daljinsku biometrijsku identifikaciju u javno dostupnim prostorima za svrhe provođenja zakona, temeljena na agregiranim podacima iz država članica na temelju godišnjih izvješća navedenih u stavku 5, koja neće uključivati osjetljive operativne podatke povezane s aktivnostima provođenja zakona.

Bez obzira na to je li sustav umjetne inteligencije stavljen na tržište ili pušten u rad neovisno o proizvodima navedenim u točkama (a) i (b), taj sustav umjetne inteligencije smatrat će se visokorizičnim ako su ispunjena oba sljedeća uvjeta:

- sustav umjetne inteligencije namijenjen je za upotrebu kao sigurnosna komponenta proizvoda, ili sam sustav umjetne inteligencije je proizvod, obuhvaćen zakonodavstvom Unije o usklađivanju navedenim u Prilogu II;
- proizvod čija je sigurnosna komponenta prema točki (a) sustav umjetne inteligencije, ili sam sustav umjetne inteligencije kao proizvod, mora proći procjenu usklađenosti treće strane, s ciljem stavljanja na tržište ili puštanja u rad tog proizvoda u skladu s zakonodavstvom Unije o usklađivanju navedenim u Prilogu II.

Pored sustava umjetne inteligencije visokog rizika spomenutih u prvom stavku, sustavi umjetne inteligencije navedeni u Prilogu III također se smatraju visokorizičnima. Međutim, sustavi umjetne inteligencije neće se smatrati visokorizičnima ako ne predstavljaju značajan rizik za zdravlje, sigurnost ili temeljna prava prirodnih osoba, uključujući neutjecanje na ishod donošenja odluka. To će biti slučaj ako je ispunjen jedan ili više sljedećih kriterija:

- Sustav umjetne inteligencije namijenjen je izvođenju uskog proceduralnog zadatka.
- Sustav umjetne inteligencije namijenjen je poboljšanju rezultata prethodno završene ljudske aktivnosti.
- Sustav umjetne inteligencije namijenjen je otkrivanju obrazaca donošenja odluka ili odstupanja od prethodnih obrazaca donošenja odluka i nije namijenjen zamjeni ili utjecaju na prethodno završenu ljudsku procjenu bez odgovarajućeg ljudskog pregleda.
- Sustav umjetne inteligencije namijenjen je izvođenju pripremnog zadatka za procjenu relevantnu za svrhu slučajeva upotrebe navedenih u Prilogu III (Europska komisija, 2024).

Unatoč prvom podstavku ovog odlomka, sustav umjetne inteligencije uvijek će se smatrati visokorizičnim ako obavlja profiliranje prirodnih osoba. Davatelj koji smatra da sustav umjetne inteligencije naveden u Prilogu III nije visokorizičan mora dokumentirati svoju procjenu prije nego što se sustav stavi na tržište ili se pusti u rad. Takav davatelj podliježe obvezi registracije utvrđenoj u članku 51(1a). Na zahtjev nacionalnih nadležnih tijela, davatelj mora pružiti dokumentaciju procjene.

Komisija će, nakon savjetovanja s Odborom za AI, najkasnije 18 mjeseci nakon stupanja na snagu ove Uredbe, pružiti smjernice koje preciziraju praktičnu primjenu ovog članka, dopunjene sveobuhvatnim popisom praktičnih primjera slučajeva upotrebe visokog i niskog rizika sustava umjetne inteligencije u skladu s uvjetima navedenim u članku 82a. Komisija je

ovlaštena usvajati delegirane akte u skladu s člankom 73 kako bi izmijenila kriterije utvrđene u točkama (a) do (d) prvog podstavka odlomka 2a. Komisija može dodati nove kriterije tim kriterijima ili ih mijenjati samo ako postoji konkretni i pouzdani dokaz o postojanju sustava umjetne inteligencije koji spadaju pod opseg Priloga III, ali ne predstavljaju značajan rizik štete za zdravlje, sigurnost i temeljna prava (Europska komisija, 2024).

Komisija će usvajati delegirane akte kako bi uklonila bilo koji od kriterija navedenih u prvom podstavku odlomka 2a ako postoji konkretni i pouzdani dokaz da je to potrebno za održavanje razine zaštite zdravlja, sigurnosti i temeljnih prava u Uniji. Svaka izmjena kriterija navedenih u točkama (a) do (d) u prvom podstavku odlomka 2a ne smije smanjiti ukupnu razinu zaštite zdravlja, sigurnosti i temeljnih prava u Uniji. Pri usvajanju delegiranih akata, Komisija će osigurati usklađenost s delegiranim aktima usvojenim u skladu s člankom 7(1) i uzeti u obzir tržišne i tehnološke razvoje (Europska komisija, 2024).

### 3.3. AI opće namjene (GPAI)

Tehnologije umjetne inteligencije opće namjene, poput ChatGPT-a, brzo mijenjaju način na koji se sustavi umjetne inteligencije izrađuju i primjenjuju. Iako se očekuje da će te tehnologije donijeti velike koristi u narednim godinama, potičući inovacije u mnogim sektorima, njihova disruptivna priroda izaziva pitanja politike vezana uz privatnost i prava intelektualnog vlasništva, odgovornost i odgovornost te zabrinutost zbog njihovog potencijala za širenje dezinformacija i pogrešnih informacija. Zakonodavci EU-a trebaju naći delikatan balans između poticanja primjene ovih tehnologija uz osiguranje adekvatnih zaštitnih mjera (Europski parlament, 2023).

Iako globalno usuglašena definicija umjetne inteligencije ne postoji, znanstvenici uglavnom dijele mišljenje da tehnički postoje dvije široke kategorije tehnologija umjetne inteligencije: 'umjetna uska inteligencija' (ANI) i 'umjetna opća inteligencija' (AGI). Tehnologije ANI, poput sustava za prepoznavanje slika i govora, također nazvane slabom AI, trenirane su na dobro označenim skupovima podataka kako bi obavljale specifične zadatke i djelovale unutar predefiniranog okruženja. Za razliku od njih, tehnologije AGI, koje se također nazivaju snažnom AI, su strojevi dizajnirani da izvode širok raspon inteligentnih zadataka, razmišljaju apstraktno i prilagođavaju se novim situacijama.

Iako je razvoj AGI prije samo nekoliko godina izgledao umjeren, brzi tehnološki proboji, uključujući korištenje tehnika velikih jezičnih modela (LLM), od tada su radikalno promijenili potencijal ovih tehnologija. Nova generacija tehnologija AGI s generativnim sposobnostima – koje se nazivaju 'umjetnom inteligencijom opće namjene' ili 'temeljnim modelima' – trenira se na širokom skupu neoznačenih podataka koji se mogu koristiti za različite zadatke s minimalnim fino podešavanjem. Ti temeljni modeli postaju dostupni razvojnim programerima niže razine putem sučelja za programiranje aplikacija (API) i otvorenog pristupa te se danas koriste kao infrastruktura od strane mnogih tvrtki za pružanje usluga krajnjim korisnicima (Europski parlament, 2023).

Alati umjetne inteligencije opće namjene sada postaju dostupni široj javnosti. U 2020. godini istraživački laboratorij OpenAI, koji je u međuvremenu sklopio komercijalno partnerstvo s Microsoftom, objavio je GPT-3, model jezika obučen na velikim internet skupovima podataka koji može obavljati širok raspon zadataka obrade prirodnog jezika, uključujući prijevod jezika, sažimanje i odgovaranje na pitanja. U 2021. OpenAI je objavio DALL-E, dubinski model učenja koji može generirati digitalne slike iz opisa prirodnog jezika. U prosincu 2022. pokrenuo je svoj chatbot ChatGPT, baziran na GPT-3 i obučen na modelima strojnog učenja koristeći podatke s interneta za generiranje bilo koje vrste teksta. U ožujku 2023. lansiran je GPT-4, najnoviji alat umjetne inteligencije opće namjene, od kojeg se očekuje da će imati još više primjena u područjima poput kreativnog pisanja, generiranja umjetnosti i programiranja (Europski parlament, 2023).

U ožujku 2023. Microsoft je lansirao novi AI-pokretani pretraživač Bing i preglednik Edge koji uključuju funkciju chata koja donosi više konteksta rezultatima pretraživanja. Također je objavio platformu GPT-4 koja omogućava tvrtkama da izgrade vlastite aplikacije, na primjer za sažimanje sadržaja velikog formata i pomoć u pisanju softvera. Google i njegova podružnica DeepMind također razvijaju alate umjetne inteligencije opće namjene; primjeri uključuju uslugu razgovorne AI, Bard. Google je u ožujku 2023. otkrio niz generativnih AI alata, dajući poslovnim subjektima i vladama mogućnost generiranja teksta, slika, koda, videozapisa, audiozapisa i izgradnje vlastitih aplikacija. Programeri koriste te 'temeljne modele' za izradu i ponudu niza novih AI usluga krajnjim korisnicima.

Alati umjetne inteligencije opće namjene imaju potencijal transformirati mnoga područja, na primjer stvaranjem novih arhitektura pretraživača ili personaliziranih terapijskih botova, ili pomažući programerima u njihovim programerskim zadacima. Prema studiji Gartnera, ulaganja u generativna AI rješenja sada vrijede više od 1,7 milijardi američkih dolara. Studija predviđa da će generativna AI u narednim godinama imati snažan utjecaj na zdravstvo, proizvodnju, automobilsku industriju, zrakoplovstvo i obranu, među ostalima. Generativna AI može se koristiti u medicinskom obrazovanju i potencijalno u kliničkom donošenju odluka ili u dizajniranju novih lijekova i materijala. Mogla bi čak postati ključni izvor informacija u zemljama u razvoju za rješavanje nedostatka stručnosti.

Ključne karakteristike identificirane u modelima umjetne inteligencije opće namjene – njihova veličina, nejasnoća i potencijal za razvijanje neočekivanih sposobnosti koje nadilaze namjere njihovih proizvođača – postavljaju mnoga pitanja. Studije su dokumentirale da veliki jezični modeli (LLMs), poput ChatGPT-a, predstavljaju etičke i društvene rizike. Mogu nepravedno diskriminirati, perpetuirati stereotipe i društvene predrasude, koristiti toksičan jezik (na primjer poticati mržnju ili nasilje), predstavljati rizik za osobne i osjetljive informacije, pružati lažne ili obmanjujuće informacije, povećati učinkovitost kampanja dezinformacija te uzrokovati niz šteta u interakciji čovjeka i računala (kao što je navođenje korisnika da precjenjuju sposobnosti AI-a i koriste ga na nesigurne načine). Unatoč pokušajima inženjera da ublaže te rizike, LLM-ovi, poput GPT-4, i dalje predstavljaju izazove za sigurnost korisnika i temeljna prava (na primjer proizvodeći uvjerljiv tekst koji je suptilno lažan ili pokazujući povećanu sposobnost za pružanje nezakonitih savjeta) te mogu generirati štetan i kriminalan sadržaj (Europski parlament, 2023).

Budući da se modeli umjetne inteligencije opće namjene treniraju skupljanjem, analizom i obradom javno dostupnih podataka s interneta, stručnjaci za privatnost naglašavaju probleme privatnosti u vezi s plagijarizmom, transparentnošću, pristankom i zakonitim osnovama za obradu podataka. Ti modeli predstavljaju izazov za obrazovne sustave i zajedničke resurse poput javnih repozitorija. Nadalje, pojava LLM-ova postavlja mnoga pitanja, uključujući kršenje prava intelektualnog vlasništva i distribuciju autorskih materijala bez dopuštenja. Neki stručnjaci upozoravaju da bi kreativnost generirana AI-om mogla značajno poremetiti kreativne industrije (na primjer u područjima grafičkog dizajna ili kompozicije glazbe). Pozivaju na poticaje koji bi ojačali inovacije i komercijalizaciju kreativnosti generirane AI-om s jedne strane, i na mjere za zaštitu vrijednosti ljudske kreativnosti s druge strane. Također se postavlja

pitanje koje bi režime odgovornosti trebalo koristiti kada sustavi umjetne inteligencije opće namjene uzrokuju štetu. Očekuje se da će ti modeli imati značajan utjecaj na tržište rada, uključujući i vrste radnih zadataka.

U tom kontekstu, stručnjaci tvrde da postoji snažna potreba za upravljanjem širenjem alata umjetne inteligencije opće namjene, s obzirom na njihov utjecaj na društvo i gospodarstvo. Također pozivaju na nadzor i praćenje LLM-ova putem mehanizama evaluacije i testiranja, naglašavajući opasnost od dopuštanja da ti alati ostanu u rukama samo nekoliko tvrtki i vlada te ističući potrebu za procjenom složenih ovisnosti između tvrtki koje razvijaju i primjenjuju alate umjetne inteligencije opće namjene. Stručnjaci za AI također pozivaju na najmanje šestomjesečnu pauzu u treniranju AI sustava moćnijih od GPT-4 (Europski parlament, 2023).

Trenutno se zakonodavci EU-a bave dugotrajnim pregovorima kako bi definirali regulatorni okvir EU-a za umjetnu inteligenciju koji bi podvrgnuo sustave umjetne inteligencije visokog rizika skupu zahtjeva i obveza unutar EU-a. Točan opseg predloženog zakona o umjetnoj inteligenciji (AI zakon) predmet je rasprave. Iako izvorni prijedlog Europske komisije nije sadržavao posebne odredbe o tehnologijama umjetne inteligencije opće namjene, Vijeće je predložilo da se one uzmu u obzir. Znanstvenici su upozorili da bi svaki pristup koji klasificira sustave AI kao visokorizične ili ne, ovisno o njihovoj namjeravanoj svrsi, stvorio prazninu za sustave opće namjene, budući da bi budući AI zakon regulirao specifične upotrebe AI aplikacije, ali ne i njezine temeljne modele.

U tom kontekstu, brojni dionici, pozvali su da se umjetna inteligencija opće namjene uključi u opseg AI zakona. Ti bi sustavi bili podložni pravnim obvezama i zahtjevima koji odgovaraju njihovim karakteristikama te sustavu nadzora sistemskog rizika sličnom onome pod Zakonom o digitalnim uslugama (DSA). AI zakon trebao bi se usredotočiti na specifične visokorizične primjene umjetne inteligencije opće namjene i uključivati obveze u vezi s transparentnošću, upravljanjem rizicima i nediskriminacijom; pravila o moderiranju sadržaja DSA (na primjer mehanizmi obavještanja i djelovanja, te pouzdani označivači) trebala bi se proširiti kako bi obuhvatila takvu umjetnu inteligenciju opće namjene.

### 3.4. Upravljanje i provedba EU Akta



Implementacija Zakona o umjetnoj inteligenciji bit će upravljana osnivanjem Ureda za AI, koji će djelovati unutar Europske komisije. Ured će nadgledati učinkovitu provedbu i poštivanje odredbi Zakona od strane pružatelja modela umjetne inteligencije opće namjene (GPAI). Osim toga, pružatelji usluga niže razine moći će podnijeti pritužbu Uredu za AI u slučajevima kada smatraju da pružatelji usluga više razine krše odredbe Zakona.

Ured za AI će imati mogućnost provođenja evaluacija modela GPAI kako bi:

1. Procijenio usklađenost s odredbama Zakona, posebno kada informacije prikupljene kroz njegove ovlasti za zahtijevanje informacija budu nedovoljne.
2. Istražio sistematske rizike, posebno nakon primanja kvalificiranog izvještaja od znanstvenog panela neovisnih stručnjaka.

Takav pristup omogućava stalni nadzor i prilagodbu praksi i politika u vezi s umjetnom inteligencijom, osiguravajući da razvoj i primjena GPAI modela budu u skladu s etičkim normama i zakonskim zahtjevima, te zaštitom prava i sigurnosti korisnika.

AI Akt je ključni zakonodavni dokument koji uspostavlja pravila za razvoj i korištenje umjetne inteligencije unutar EU, s naglaskom na zaštitu ljudskih prava i osiguravanje sigurnog i etičkog korištenja AI tehnologija. AI Akt, formalno nazvan Uredba (EU) 2024/1689 Europskog parlamenta i Vijeća od 13. lipnja 2024., donesen je s ciljem uspostavljanja usklađenih pravila o umjetnoj inteligenciji unutar EU. Ova uredba također donosi izmjene niza prethodnih uredbi i direktiva kako bi se stvorio koherentan pravni okvir za razvoj, stavljanje na tržište i upotrebu AI sustava. Uredba stupa na snagu dvadesetog dana nakon objave u Službenom listu EU, a njene odredbe postaju obvezne za sve države članice. To znači da će sve države članice morati prilagoditi svoje nacionalne zakone kako bi bili u skladu s ovom uredbom u vrlo kratkom vremenskom roku.

Glavni cilj Akta je uspostavljanje ujednačenog pravnog okvira unutar EU za razvoj, distribuciju i korištenje AI sustava. Ovaj okvir je usklađen s temeljnim vrijednostima EU-a, poput zaštite ljudskih prava, demokracije, vladavine prava i zaštite okoliša. Regulacija također teži jačanju povjerenja u AI sustave, osiguravajući da se oni koriste na siguran i etički način, te da se spriječe potencijalne štete koje mogu nastati nepravilnim ili zlonamjernim korištenjem AI tehnologija. U prologu Akta ističe se potreba za ovakvom regulacijom zbog brzog razvoja AI tehnologija koje imaju potencijalno širok utjecaj na društvo i gospodarstvo. Akt također spominje

prethodne pokušaje pojedinih država članica da donesu nacionalne regulative za AI, što bi moglo dovesti do fragmentacije unutar jedinstvenog tržišta. Zbog toga je donesena ova uredba kako bi se osiguralo da AI tehnologije budu sigurne, pouzdane i u skladu s pravnim okvirom EU-a.

## 4. ZAKLJUČAK

Akt o umjetnoj inteligenciji EU-a postavlja se kao ključna točka u reguliranju i upravljanju AI tehnologijama, posebice onima visokog rizika i opće namjene, kako bi se osiguralo njihovo etičko i sigurno korištenje. Kako se tehnologija umjetne inteligencije razvija eksponencijalnom brzinom, ključno je postavljanje čvrstih pravila koja će pratiti njezinu implementaciju i integraciju u društvo.

Regulacija AI-a nije samo pitanje tehnološke sigurnosti, već i očuvanja osnovnih ljudskih prava i sloboda. Iako AI nudi iznimne mogućnosti za poboljšanje životnih i radnih uvjeta, postoje ozbiljni rizici, poput zloupotrebe, pristranosti i gubitka privatnosti, koji zahtijevaju pažljivo reguliranje. EU stoga nastoji postaviti globalne standarde transparentnosti, sigurnosti i odgovornosti kroz svoj AI zakon. Dugoročno, AI zakon bi trebao omogućiti harmonizaciju tehnoloških praksi unutar granica EU-a, sprječavajući fragmentaciju koja bi mogla nastati zbog različitih nacionalnih pravila. Ovo je posebno važno za očuvanje jedinstvenog tržišta i poticanje inovacija unutar bloka, dok se istovremeno osigurava da AI tehnologije ne postanu alati za nadzor ili sredstva za diskriminaciju.

Za pružatelje AI usluga, ovaj zakon predstavlja priliku da poboljšaju svoje procese i osiguraju da njihovi proizvodi zadovoljavaju visoke standarde etičnosti i sigurnosti. S druge strane, za korisnike i potrošače, AI zakon donosi veću sigurnost i zaštitu, dajući im veću kontrolu nad svojim digitalnim otiscima i podacima. Zakon o umjetnoj inteligenciji treba gledati ne samo kao regulatorni okvir, već i kao kritičnu komponentu u oblikovanju budućnosti umjetne inteligencije u skladu s etičkim standardima i ljudskim pravima. Ovo je ključno za osiguravanje

da AI služi javnom dobru, promiče socijalnu i ekonomsku dobrobit, i da tehnološki napredak ide ruku pod ruku s etičkim razvojem.

## POPIS LITERATURE

- 1) AAAI. (2022). A Taxonomy for Artificial Intelligence. Association for the Advancement of Artificial Intelligence. Dostupno na: <https://aaai.org/> Pristup 28.04.2024.
- 2) Agrawal, A., Gans, J. S., Goldfarb, A. (2019). Artificial intelligence: the ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives*, 33(2), 31-50.
- 3) AI Impacts (2021). AI Timeline Surveys. Dostupno na: <https://aiimpacts.org/ai-timeline-surveys/> Pristup 29.04.2024.
- 4) Ameen, N., Sharma, G. D., Tarba, S., Rao, A., Chopra, R. (2022). Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology & Marketing*, 39(9), 1802-1825.
- 5) Ariffin, A. S., Maavak, M., Dolah, R., Muhtazaruddin, M. N. (2023). Formulation of AI Governance and ethics framework to support the implementation of responsible AI for Malaysia. *Res Militaris*, 13(3), 2491-2516.
- 6) Bostrom, N. (2016). The control problem, Excerpts from *superintelligence: Paths, dangers, strategies*. *Science Fiction and Philosophy: From Time Travel to Superintelligence*, 308-330.
- 7) Boucher, P. (2020). *Artificial intelligence: How does it work, why does it matter, and what can we do about it?* Brussels: European parliament.
- 8) Churchland, P. M. (1996). *The engine of reason, the seat of the soul: A philosophical journey into the brain*. MIT Press.

- 9) Dagnaw, G. (2020). Artificial intelligence towards future industrial opportunities and challenges. Kennesaw State University.
- 10) Davenport, T. H., Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108-116.
- 11) Europska komisija (2024). EU AI Act - EU Artificial Intelligence Act. Dostupno na: <https://www.euaiact.com/> Pristup 03.05.2024.
- 12) Europska komisija. (2021). Regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Europa.eu. Dostupno na: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206> Pristup 02.05.2024.
- 13) Europski parlament (2023). AT A GLANCE Digital issues in focus. Dostupno na: [https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS\\_ATA\(2023\)745708\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/745708/EPRS_ATA(2023)745708_EN.pdf) Pristup 04.05.2024.
- 14) Gillham, J., Rimmington, L., Dance, H., Verweij, G., Rao, A., Roberts, K. B., Paich, M. (2018). The macroeconomic impact of artificial intelligence. PwC Report-PricewaterhouseCoopers.-2018.
- 15) Goertzel, B. (2007). Artificial general intelligence (Vol. 2, p. 1). C. Pennachin (Ed.). New York: Springer.
- 16) Haenlein, M., Kaplan, A., Tan, C. W., Zhang, P. (2019). Artificial intelligence (AI) and management analytics. *Journal of Management Analytics*, 6(4), 341-343.
- 17) Jajal, D., T. (2021). Distinguishing between Narrow AI, General AI and Super AI. Dostupno na: <https://medium.com/mapping-out-2050/distinguishing-between-narrow-ai-general-ai-and-super-ai-a4bc44172e22> Pristup 28.04.2024.
- 18) Kamlaris, A., Kartakoullis, A., Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, 143, 23-37.
- 19) Kurzweil, R. (2017). AI Will Not Displace Humans, It's Going to Enhance Us. Dostupno na: <https://futurism.com/ray-kurzweil-ai-displace-humans-going-enhance> Pristup 29.04.2024.
- 20) Lanfranco, A. R., Castellanos, A. E., Desai, J. P., Meyers, W. C. (2004). Robotic surgery: a current perspective. *Annals of Surgery*, 239(1), 14.
- 21) LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

- 22) Poole, D. L., Mackworth, A. K. (2010). *Artificial Intelligence: foundations of computational agents*. Cambridge University Press.
- 23) Russell, S., Norvig, P. (1996). *Artificial Intelligence: A Modern Approach*, Englewood Cliffs. NJ: Prentice-Hall.
- 24) Samoili, S., Lopez Cobo, M., Gomez Gutierrez, E., De Prato, G., Martinez-Plumed, F., Delipetrev, B. (2020). *AI WATCH: Defining artificial intelligence*. Publications Office of the European Union, Luxembourg.
- 25) Surden, H. (2014). Machine learning and law. *Washington Law Review*, 89, 87.
- 26) Yannakakis, G. N., Togelius, J. (2018). *Artificial intelligence and games* (Vol. 2, pp. 2475-1502). New York: Springer.