FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

PRAVILNO RAVNANJE V PRIHODNJE

UMETNA INTELIGENCA IN TEMELJNE PRAVICE

POROČILO

REPUBLIKA SLOVENIJA
ZAGOVORNIK NAČELA ENAKOSTI

Republika Slovenija
Zagovornik načela enakosti
Železna cesta 16 – 1000 Ljubljana – Slovenija
T: + 386 (0)1 4735 531
E: **gp@zagovornik-rs.si**
zagovornik.si

Prevod iz angleškega jezika (Getting the future right – Artificial intelligence and fundamental rights in the EU) je zagotovil Zagovornika načela enakosti. Agencija Evropske unije za temeljne pravice (FRA) ne odgovarja za pravilnost prevoda.

Luxembourg: Urad za publikacije Evropske unije, 2024

Print	ISBN 978-92-9489-414-4	doi:10.2811/739026	TK-03-20-119-SL-C
PDF	ISBN 978-92-9489-415-1	doi:10.2811/842248	TK-03-20-119-SL-N

Predgovor

Ste vedeli, da umetna inteligenca že igra vlogo pri odločanju, kolikšno nadomestilo za brezposelnost bo nekdo dobil, kje je verjetno, da bo prišlo do vloma, ali pri določeni osebi obstaja tveganje, da zboli za rakom, ali komu se bo prikazal privlačen oglas za nizke hipotekarne obrestne mere?

O umetni inteligenci (UI) govorimo, ko stroji počnejo stvari, ki so jih nekoč lahko počeli samo ljudje. V današnjem času je umetna inteligenca bolj prisotna v našem življenju, kot se zavedamo – in njena uporaba še vedno narašča. Možnosti se zdijo neskončne. Toda kako lahko pri uporabi umetne inteligence v celoti spoštujemo standarde temeljnih pravic?

V tem poročilu so predstavljeni konkretni primeri, kako podjetja in javne uprave v Evropski uniji (EU) uporabljajo ali poskušajo uporabljati umetno inteligenco. Obravnavane so možne posledice za temeljne pravice, prav tako je prikazano, ali uporabniki umetne inteligence upoštevajo pravice in na kakšen način.

FRA je opravila razgovore z več kot sto uslužbenci javne uprave, zaposlenimi v zasebnih podjetjih ter različnimi strokovnjaki – vključno z organi za pregled in nadzor, nevladnimi organizacijami in odvetniki – ki se pri svojem delu na različne načine srečujejo s področjem umetne inteligence.

Na podlagi teh razgovorov je bila v poročilu izdelana analiza, kako so pri uporabi in razvoju aplikacij z umetno inteligenco upoštevane temeljne pravice. Ta se osredotoča na štiri temeljna področja – socialne prejemke, napovedno policijsko delo, zdravstvene storitve in ciljno usmerjeno oglaševanje. Načini uporabe umetne inteligence se razlikujejo glede na to, kako kompleksni so, v kolikšni meri je vključena avtomatizacija, kakšen je njihov potencialni vpliv na ljudi in kako široko se uporabljajo.

Izsledki analize poudarjajo, da je na tem področju za vse nas še veliko dela.

Eden od načinov za krepitev varstva pravic je zagotoviti, da so ljudem v primerih, ko gre kaj narobe, na voljo ustrezna pravna sredstva. A pogoj je, da morajo najprej vedeti, da se v njihovem primeru uporablja umetna inteligenca. To hkrati pomeni, da morajo biti organizacije, ki uporabljajo umetno inteligenco, sposobne pojasniti svoje sisteme umetne inteligence in razložiti, kako na njihovi podlagi sprejemajo odločitve.

Vendar so lahko zadevni sistemi resnično zapleteni. Tako tisti, ki uporabljajo sisteme umetne inteligence, kot tisti, ki so odgovorni za pravno urejanje njihove uporabe, priznavajo, da teh sistemov včasih v celoti ne razumejo. Zato je ključnega pomena zaposlovanje kadrov z ustreznim tehničnim znanjem.

Manjka namreč tudi zavedanje o morebitnih posledicah za pravice. Večina se sicer zaveda, da na področju varstva podatkov obstajajo določeni pomisleki, nekateri omenjajo tudi varstvo pred diskriminacijo. Nekoliko manjša je seznanjenost z dejstvom, da so ogrožene tudi druge pravice, med drugim denimo pravica do človekovega dostojanstva, dostop do sodnega varstva in varstvo potrošnikov. Ni presenetljivo, da se razvijalci pri presoji možnih vplivov sistemov umetne inteligence osredotočajo predvsem na tehnične vidike.

Pri spopadanju s temi izzivi moramo spodbuditi tiste, ki se ukvarjajo z varstvom človekovih pravic, in tiste, ki se ukvarjajo z umetno inteligenco, k medsebojnemu sodelovanju in izmenjavi prepotrebnega znanja, zlasti s področja tehnologije in pravic.

Tisti, ki razvijajo in uporabljajo umetno inteligenco, morajo imeti na voljo tudi ustrezna orodja za celovito oceno njenih posledic za temeljne pravice, od katerih mnoge morda niso na prvi pogled opazne. Dostopne ocene učinkov na temeljne pravice lahko spodbudijo tak razmislek in pomagajo zagotoviti, da je uporaba umetne inteligence skladna s pravnimi standardi.

Razgovori kažejo, da je uporaba umetne inteligence v EU še vedno v povojih, čeprav narašča. Toda tehnologija se razvija hitreje od prava. V tem trenutku moramo izkoristiti priložnost in zagotoviti, da bo prihodnji regulativni okvir EU za umetno inteligenco trdno zasidran v spoštovanju človekovih in temeljnih pravic.

Upamo, da bodo empirični dokazi in analiza, predstavljeni v tem poročilu, spodbudili oblikovalce politike, da sprejmejo ta izziv.

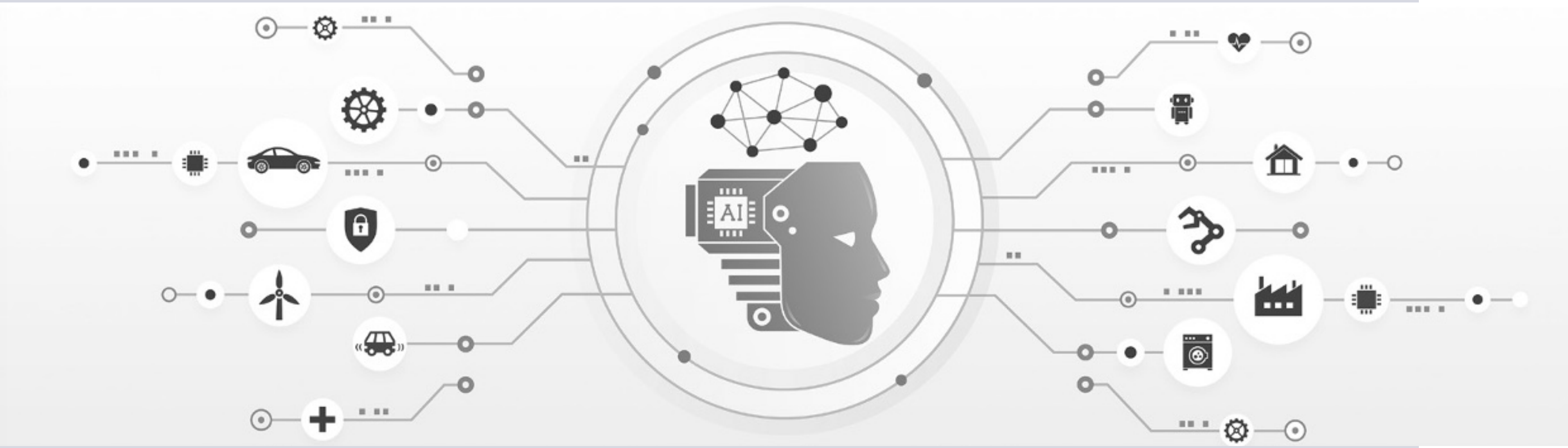
Sirpa Rautio
direktorica

Kazalo

Slike

Ključne ugotovitve in mnenja FRA

Nove tehnologije so temeljito spremenile način organiziranja naših življenj in življenje nasploh. Nove podatkovno vodene tehnologije so spodbudile zlasti razvoj umetne inteligence, vključno z večjo avtomatizacijo nalog, ki jih običajno opravljajo ljudje. Zdravstvena kriza zaradi covida-19 je spodbudila sprejemanje umetne inteligence in izmenjavo podatkov, kar je ustvarilo nove priložnosti, a tudi izzive in grožnje za človekove in temeljne pravice.



Mediji, civilna družba, akademiki, organi za človekove pravice in oblikovalci politik so veliko pozornosti namenili razvoju na področju umetne inteligence. Večina te pozornosti se osredotoča na njegov potencial za podporo gospodarski rasti. Manj pozornosti je bilo namenjene vprašanju, kako lahko razne tehnologije vplivajo na temeljne pravice. Za zdaj še nimamo veliko empiričnih dokazov o širokem spektru pravic, ki zadevajo umetno inteligenco, ali o zaščitnih ukrepih, potrebnih za zagotovitev, da je uporaba umetne inteligence v praksi v skladu s temeljnimi pravicami.

Evropska komisija je 19. februarja 2020 objavila Belo knjigo o umetni inteligenci – evropski pristop k odličnosti in zaupanju. V njej so opisana glavna načela prihodnjega regulativnega okvira EU za umetno inteligenco v Evropi. Bela knjiga ugotavlja, da je ključno, da tak okvir temelji na temeljnih vrednotah EU, vključno s spoštovanjem človekovih pravic – člen 2 Pogodbe o Evropski uniji (PEU).

To poročilo o umetni inteligenci in temeljnih pravicah podpira ta cilj z analizo posledic za temeljne pravice pri uporabi umetne inteligence. Na podlagi konkretnih primerov uporabe umetne inteligence na izbranih področjih se osredotoča na razmere na terenu v smislu izzivov in priložnosti na področju temeljnih pravic pri uporabi umetne inteligence.

Pravni okvir

Splošni okvir za temeljne pravice*, ki se uporablja za uporabo umetne inteligence v EU, sestavljata Listina EU o temeljnih pravicah (v nadaljevanju: Listina) in Evropska konvencija o človekovih pravicah.

Pomembni so še številni drugi instrumenti Sveta Evrope in mednarodni instrumenti s področja človekovih pravic. Med njimi so Splošna deklaracija o človekovih pravicah iz leta 1948 in glavne konvencije Organizacije združenih narodov (OZN) o človekovih pravicah**.

Poleg tega sekundarno pravo EU za posamezne sektorje, zlasti pravni red EU o varstvu podatkov in zakonodaja EU na področju varstva pred diskriminacijo, pomaga varovati temeljne pravice v okviru umetne inteligence. Ne nazadnje se uporabljajo tudi nacionalne zakonodaje držav članic EU.

* *Za več informacij glej FRA (2012),* **Bringing rights to life: The fundamental rights landscape of the European Union** *(Uresničevanje pravic: pregled stanja temeljnih pravic v Evropski uniji), Luxembourg, Urad za publikacije Evropske unije.*

** *Te glavne konvencije vključujejo: Mednarodni pakt o državljanskih in političnih pravicah iz leta 1966, Mednarodni pakt o ekonomskih, socialnih in kulturnih pravicah iz leta 1966, Mednarodno konvencijo o odpravi vseh oblik rasne diskriminacije iz leta 1965, Konvencijo o odpravi vseh oblik diskriminacije žensk iz leta 1979, Konvencijo proti mučenju iz leta 1984, Konvencijo o otrokovih pravicah iz leta 1989, Konvencijo o pravicah invalidov iz leta 2006 in Mednarodno konvencijo o zaščiti vseh oseb pred prisilnim izginotjem iz leta 2006.*

Za več informacij o splošnem okviru mednarodnega prava o človekovih pravicah, vključno z mehanizmi izvrševanja, glej npr. De Schutter, O. (2015), International Human Rights Law: Cases, Materials, Commentary *(Mednarodno pravo človekovih pravic: primeri, gradivo, komentar), Cambridge, Cambridge University Press, 2. izdaja.*

Poročilo temelji na 91 razgovorih z uradniki v javni upravi in zaposlenimi v zasebnih podjetjih v izbranih državah članicah EU. Postavljena so jim bila vprašanja o njihovi uporabi umetne inteligence, njihovem zavedanju o vprašanjih, povezanih s temeljnimi pravicami, in praksah v zvezi z ocenjevanjem in zmanjševanjem tveganj, povezanih z uporabo umetne inteligence.

Poleg tega je bilo opravljenih deset razgovorov s strokovnjaki, ki se na različne načine ukvarjajo s potencialnimi izzivi umetne inteligence na področju temeljnih pravic. Ta skupina je vključevala javne organe (kot so nadzorni in pregledni organi), nevladne organizacije in odvetnike.

VARSTVO TEMELJNIH PRAVIC – PODROČJE UPORABE, OCENE UČINKA IN ODGOVORNOST

Upoštevanje polnega obsega temeljnih pravic v zvezi z umetno inteligenco

Uporaba sistemov umetne inteligence vključuje širok spekter temeljnih pravic ne glede na področje uporabe. Te med drugim vključujejo zasebnost, varstvo podatkov, varstvo pred diskriminacijo in dostop do pravnega varstva.

Listina EU o temeljnih pravicah (v nadaljevanju: Listina) je postala pravno zavezujoča decembra 2009 in ima enako pravno veljavnost kot Pogodbi EU. V samo enem besedilu združuje državljanske, politične, ekonomske in socialne pravice. V skladu s členom 51(1) Listine morajo institucije, organi, uradi in agencije Unije spoštovati vse pravice iz Listine. Države članice EU jih morajo spoštovati, ko izvajajo pravo Unije. Kot za vsako drugo področje to velja tudi za umetno inteligenco.

Delo na terenu v okviru te raziskave kaže, da se v okviru umetne inteligence uporablja veliko različnih sistemov. Analizirane tehnologije vključujejo različne ravni avtomatizacije in kompleksnosti. Razlikujejo se tudi po obsegu in morebitnem vplivu na ljudi.

Ugotovitve FRA kažejo, da uporaba sistemov umetne inteligence vključuje širok spekter temeljnih pravic ne glede na področje uporabe. Te med drugim vključujejo zasebnost, varstvo podatkov, varstvo pred diskriminacijo in dostop do pravnega varstva. Vendar je pri obravnavi učinka umetne inteligence v zvezi s temeljnimi pravicami iz razgovorov razvidno, da je področje uporabe pogosto omejeno na posebne pravice.

Pri uporabi umetne inteligence je treba upoštevati širši spekter pravic, odvisno od tehnologije in področja uporabe. Poleg pravic v zvezi z zasebnostjo in varstvom podatkov, enakostjo in varstvom pred diskriminacijo ter dostopom do pravnega varstva bi lahko upoštevali tudi druge pravice. Med njimi so na primer človekovo dostojanstvo, pravica do socialne varnosti in socialne pomoči, pravica do dobrega upravljanja (večinoma pomembna za javni sektor) in varstvo potrošnikov (zlasti pomembno za podjetja). Glede na kontekst uporabe umetne inteligence je treba upoštevati vse druge pravice, zaščitene z Listino.

MNENJE FRA ŠT. 1

Pri uvajanju novih politik in sprejemanju nove zakonodaje o umetni inteligenci morajo zakonodajalec EU in države članice v okviru področja uporabe prava EU zagotoviti, da se upošteva celoten spekter temeljnih pravic, kot so določene v Listini in Pogodbah EU. Zadevne politike in zakone morajo spremljati specifični zaščitni ukrepi za temeljne pravice.

Pri tem bi se morale EU in njene države članice opirati na trdne dokaze o vplivu umetne inteligence na temeljne pravice, da bi zagotovile, da se pri omejevanju nekaterih temeljnih pravic spoštujeta načeli nujnosti in sorazmernosti.

Z zakonom je treba zagotoviti ustrezne zaščitne ukrepe za učinkovito zaščito pred samovoljnim poseganjem v temeljne pravice ter za zagotovitev pravne varnosti razvijalcem in uporabnikom umetne inteligence. Prostovoljne sheme za opazovanje in varstvo temeljnih pravic pri razvoju in uporabi umetne inteligence lahko dodatno pomagajo ublažiti kršitve pravic. V skladu z minimalnimi zahtevami pravne jasnosti – kot osnovnega načela pravne države in temeljnega pogoja za zagotavljanje temeljnih pravic – mora zakonodajalec pri opredelitvi področja uporabe katerega koli takega zakona o umetni inteligenci ravnati skrbno.

Glede na raznolikost tehnologije, ki je zajeta v izraz umetna inteligenca, in pomanjkanje znanja o celotnem obsegu njenega potencialnega učinka na temeljne pravice bo morda treba redno ocenjevati pravno opredelitev izrazov, povezanih z umetno inteligenco.

MNENJE FRA ŠT. 2

Zakonodajalec EU bi moral razmisliti o uvedbi obveznih ocen učinka, ki zajemajo celoten nabor temeljnih pravic. Te ocene bi morale zajeti zasebni in javni sektor ter se izvesti pred uporabo katerega koli sistema umetne inteligence. V ocenah učinka bi bilo treba upoštevati različno naravo in obseg tehnologij umetne inteligence, vključno z ravnjo avtomatizacije in kompleksnosti, ter morebitno škodo. Vključevati bi morale osnovne zahteve za pregled, ki se lahko uporabijo tudi za ozaveščanje o potencialnih posledicah za temeljne pravice.

Ocene učinka bi morale temeljiti na uveljavljeni dobri praksi z drugih področij, njihovo izvajanje pa bi se moralo po potrebi med uvajanjem redno ponavljati. Te ocene bi se morale izvajati pregledno. Njihovi rezultati in priporočila bi morali biti čim bolj javno dostopni. V smislu podpore postopku ocene učinka bi bilo treba od podjetij in javne uprave zahtevati, da zberejo informacije, potrebne za temeljito oceno morebitnega učinka v zvezi s temeljnimi pravicami.

EU in države članice bi morale razmisliti o ciljno usmerjenih ukrepih v podporo tistim, ki razvijajo, uporabljajo ali načrtujejo uporabo sistemov umetne inteligence, da bi zagotovile učinkovito izpolnjevanje svojih obveznosti glede ocen učinka v zvezi s temeljnimi pravicami. Takšni ukrepi bi lahko vključevali financiranje, smernice, usposabljanje ali ozaveščanje. Usmerjeni bi morali biti predvsem, vendar ne izključno, v zasebni sektor.

EU in države članice bi morale razmisliti o uporabi že razvitih orodij, kot so kontrolni seznami ali orodja za samoocenjevanje, ki so bila razvita na evropski in mednarodni ravni. Med njimi so tista, ki jih je razvila skupina Evropske komisije na visoki ravni za umetno inteligenco.

Uporaba učinkovitih ocen učinka za preprečevanje negativnih učinkov

Predhodne ocene učinka se osredotočajo predvsem na tehnična vprašanja. Redko obravnavajo potencialne učinke na temeljne pravice. Razlog za to je pomanjkanje znanja o tem, kako umetna inteligenca vpliva na te pravice.

Uporaba sistemov umetne inteligence vključuje širok spekter temeljnih pravic ne glede na področje uporabe. V skladu s členom 51(1) Listine morajo države članice EU pri izvajanju prava Unije spoštovati vse pravice iz Listine. V skladu z veljavnimi mednarodnimi standardi, zlasti z vodilnimi načeli Združenih narodov o podjetništvu in človekovih pravicah, bi morala podjetja vzpostaviti „postopek v zvezi s spoštovanjem človekovih pravic, s katerim prepoznavajo, preprečujejo in zmanjšajo svoje negativne vplive na človekove pravice ter o njih poročajo“ (načeli 15 in 17). To velja ne glede na njihovo velikost in sektor, zajema pa tudi podjetja, ki se ukvarjajo z umetno inteligenco.

EU je v skladu s svojimi zavezami glede vodilnih načel OZN sprejela številne zakonodajne akte, ki obravnavajo sektorske instrumente, zlasti v okviru obveznosti v zvezi s potrebno skrbnostjo za človekove pravice. Trenutno potekajo razprave o predlogu nove sekundarne zakonodaje EU. Takšna zakonodaja bi od podjetij zahtevala, da izvajajo potrebno skrbnost v zvezi s potencialnimi učinki svojih dejavnosti in dobavnih verig na človekove pravice in okolje. Takšna zakonodaja bi bila verjetno medsektorska in bi določala sankcije za neskladnost, ki bi morale zajemati uporabo umetne inteligence. Glej nedavno poročilo FRA o **podjetništvu in človekovih pravicah – dostop do pravnih sredstev**, ki poziva k izboljšanim horizontalnim pravilom o skrbnosti na področju človekovih pravic za podjetja s sedežem v EU.

Ocene učinka so pomembno orodje za podjetja in javno upravo za ublažitev morebitnega negativnega učinka njihovih dejavnosti na temeljne pravice. Zakonodaja EU v posameznih sektorjih zahteva nekatere oblike ocen učinka, kot so ocene učinka v zvezi z varstvom podatkov v skladu s splošno uredbo o varstvu podatkov (GDPR). Številni anketiranci so poročali, da je bila ocena učinka v zvezi z varstvom podatkov izvedena, kot to zahteva zakon. Vendar so bile te ocene opravljene v različnih oblikah. Poleg tega se predhodne ocene, kadar se izvajajo, osredotočajo predvsem na tehnične vidike. Redko obravnavajo morebitne učinke na temeljne pravice. Nekateri anketiranci menijo, da se ocene učinka v zvezi s temeljnimi pravicami ne izvajajo, kadar sistem umetne inteligence nima negativnega vpliva na temeljne pravice ali ni videti, da bi ga lahko imel.

Raziskava je pokazala, da je znanje anketirancev o temeljnih pravicah – razen varstva podatkov in deloma varstva pred diskriminacijo – šibko. Vendar večina priznava, da uporaba umetne inteligence vpliva na temeljne pravice.

Nekateri anketiranci navajajo, da njihovi sistemi ne vplivajo na temeljne pravice, kar je deloma povezano z nalogami, za katere se uporabljajo sistemi umetne inteligence.

Vsi anketiranci poznajo vprašanja v zvezi z varstvom podatkov. Večina vprašanih se zaveda tudi, da bi lahko bila diskriminacija na splošno težava pri uporabi umetne inteligence. Vendar številni anketiranci še vedno ne razumejo natančnega pomena in uporabnosti pravic, povezanih z varstvom podatkov in varstvom pred diskriminacijo.

Izsledki raziskav kažejo razlike med zasebnim in javnim sektorjem. Anketiranci iz zasebnega sektorja se pogosto manj zavedajo širšega nabora temeljnih pravic, ki bi lahko bile prizadete. Zasebni sektor pozna vprašanja v zvezi z varstvom podatkov. Vendar predstavniki podjetij, ki uporabljajo umetno inteligenco, slabše poznajo druge pravice, kot je varstvo pred diskriminacijo ali dostop do pravic, povezanih s pravosodjem. Nekateri so se v celoti zavedali morebitnih težav, drugi pa so menili, da so za preverjanje vprašanj v zvezi s temeljnimi pravicami odgovorne njihove stranke.

MNENJE FRA ŠT. 3

EU in države članice bi morale zagotoviti, da so vzpostavljeni učinkoviti sistemi odgovornosti za spremljanje in po potrebi učinkovito obravnavanje morebitnih negativnih učinkov sistemov umetne inteligence na temeljne pravice. Poleg ocene učinka v zvezi s temeljnimi pravicami (glej mnenje FRA št. 2) bi morale razmisliti o uvedbi posebnih zaščitnih ukrepov za zagotovitev učinkovitega sistema odgovornosti. To bi lahko vključevalo zakonsko zahtevo, da se zagotovi dovolj informacij za oceno učinka sistemov umetne inteligence na temeljne pravice. To bi pristojnim organom omogočilo zunanje spremljanje in nadzor nad človekovimi pravicami.

EU in države članice bi morale tudi bolje izkoristiti že vzpostavljene strokovne strukture za nadzor, da bi zaščitile temeljne pravice pri uporabi umetne inteligence. Med njimi so organi za varstvo podatkov, organi za enakost, nacionalne institucije za človekove pravice, varuhi človekovih pravic in organi za varstvo potrošnikov.

Dodatna sredstva bi bilo treba nameniti vzpostavitvi učinkovitih sistemov odgovornosti z izpopolnjevanjem in zagotavljanjem raznolikosti osebja, ki dela za nadzorne organe. To bi jim omogočilo obravnavo zapletenih vprašanj, povezanih z razvojem in uporabo umetne inteligence.

Podobno bi morali imeti ustrezni organi zadostna sredstva, pooblastila in – kar je še pomembnejše – strokovno znanje za preprečevanje in ocenjevanje kršitev temeljnih pravic ter učinkovito podporo tistim, na temeljne pravice katerih vpliva umetna inteligenca.

Olajšanje sodelovanja med ustreznimi organi na nacionalni in evropski ravni lahko pripomore k izmenjavi strokovnega znanja in izkušenj. Pomaga lahko tudi sodelovanje z drugimi subjekti z ustreznim strokovnim znanjem, kot so specializirane organizacije civilne družbe. Države članice bi morale pri izvajanju takšnih ukrepov na nacionalni ravni razmisliti o uporabi razpoložljivih mehanizmov financiranja EU.

Zagotavljanje učinkovitega nadzora in splošne odgovornosti

Podjetja in javne uprave, ki razvijajo in uporabljajo umetno inteligenco, so v stiku z raznimi organi, odgovornimi za nadzor sistemov, povezanih z umetno inteligenco, v okviru svojih pristojnosti in sektorjev. Ti organi vključujejo organe za varstvo podatkov. Vendar tistim, ki uporabljajo umetno inteligenco, ni vedno povsem jasno, kateri organi so odgovorni za nadzor sistemov umetne inteligence.

V skladu z uveljavljenimi mednarodnimi standardi na področju človekovih pravic – na primer s členom 1 Evropske konvencije o človekovih pravicah (EKČP) in členom 51 Listine – morajo države varovati pravice in svoboščine ljudi. Za učinkovito izvajanje morajo države med drugim vzpostaviti učinkovite mehanizme spremljanja in izvrševanja. To velja tudi za umetno inteligenco.

Ugotovitve na ravni spremljanja kažejo, da imajo pomembno vlogo specializirani organi, ustanovljeni v določenih sektorjih, ki so v okviru svojih pristojnosti odgovorni tudi za nadzor umetne inteligence. Ti vključujejo na primer nadzor na področju bančništva ali organe za varstvo podatkov. Razni taki organi bi lahko bili pomembni za nadzor umetne inteligence z vidika temeljnih pravic. Vendar odgovornosti organov v zvezi z nadzorom umetne inteligence mnogim od vprašanih iz zasebnega in javnega sektorja ostajajo nejasne.

Uporaba umetne inteligence s strani javnih uprav se včasih revidira v okviru njihovih rednih revizij. Zasebna podjetja v posameznih sektorjih imajo tudi specializirane nadzorne organe, na primer na področju zdravstvenih ali finančnih storitev. Ti preverjajo tudi uporabo umetne inteligence in sorodnih tehnologij, na primer v okviru svojih certifikacijskih shem. Anketiranci iz zasebnega sektorja so izrazili željo po organih, ki bi lahko zagotavljali strokovne nasvete o možnostih in zakonitosti morebitnih uporab umetne inteligence.

V EU je dobro razvit sklop neodvisnih organov, ki so odgovorni za varstvo in spodbujanje temeljnih pravic. Med njimi so organi za varstvo podatkov, organi za enakost, nacionalne institucije za človekove pravice in varuhi človekovih pravic. Raziskava je pokazala, da so se tisti, ki uporabljajo umetno inteligenco ali jo nameravajo uporabljati, glede svoje uporabe umetne inteligence pogosto obrnili na razne organe, kot so organi za varstvo potrošnikov.

Da bi poiskali smernice, informacije ali odobritev v zvezi z obdelavo osebnih podatkov, so se uporabniki umetne inteligence najpogosteje obrnili na organe za varstvo podatkov. Strokovnjaki, s katerimi so bili opravljeni razgovori, poudarjajo pomen organov za varstvo podatkov za nadzor sistemov umetne inteligence v zvezi z uporabo osebnih podatkov. Vendar ugotavljajo tudi, da imajo organi za varstvo podatkov za to nalogo premalo sredstev in nimajo posebnega strokovnega znanja o vprašanjih umetne inteligence.

Strokovnjaki, vključno s tistimi, ki delajo za nadzorne organe, kot so organi za enakost in organi za varstvo podatkov, se strinjajo, da je treba okrepiti strokovno znanje obstoječih nadzornih organov, da se jim omogoči učinkovit nadzor nad vprašanji, povezanimi z umetno inteligenco. Po mnenju strokovnjakov je to lahko izziv, saj so viri teh organov že izčrpani. Poudarili so tudi pomembno vlogo ustreznih organizacij civilne družbe, specializiranih za področje tehnologije, digitalnih pravic in algoritmov. Te lahko okrepijo odgovornost pri uporabi sistemov umetne inteligence.

NEDISKRIMINACIJA, VARSTVO PODATKOV IN DOSTOP DO PRAVNEGA VARSTVA: TRI HORIZONTALNE TEME

Raziskave kažejo, da uporaba umetne inteligence vpliva na različne temeljne pravice. Poleg specifičnih vidikov, povezanih z okoliščinami, ki različno vplivajo na različne pravice, so teme temeljnih pravic, ki so se pojavile v raziskavi in se večkrat uporabljajo za večino primerov umetne inteligence, med drugim potreba po zagotavljanju nediskriminatorne uporabe umetne inteligence (pravica do nediskriminacije), zahteva po zakoniti obdelavi podatkov (pravica do varstva osebnih podatkov) ter možnost pritožbe zoper odločitve, ki temeljijo na umetni inteligenci, in uporabe pravnih sredstev (pravica do učinkovitega pravnega sredstva in poštenega sojenja).

Glavni temeljni pravici, izpostavljeni v razgovorih, sta varstvo podatkov in pravica do nediskriminacije. Poleg tega so se večkrat pojavili učinkoviti načini pritožbe glede uporabe umetne inteligence, povezani s pravico do poštenega sojenja in učinkovitega pravnega sredstva. Naslednja tri mnenja FRA, ki odražajo te ugotovitve, bi bilo treba brati skupaj z drugimi mnenji, ki pozivajo k celovitejšemu priznavanju vseh temeljnih pravic, na katere vpliva umetna inteligenca, in odzivanju nanje.

MNENJE FRA ŠT. 4

Države članice EU bi morale razmisliti o tem, da bi podjetja in javno upravo spodbudile k oceni potencialno diskriminatornih rezultatov pri uporabi sistemov umetne inteligence.

Evropska komisija in države članice bi morale razmisliti o zagotovitvi sredstev za ciljno usmerjene raziskave o potencialno diskriminatornih učinkih uporabe umetne inteligence in algoritmov. Takšnim raziskavam bi koristila prilagoditev uveljavljenih raziskovalnih metod iz družbenih ved, ki se uporabljajo za odkrivanje morebitne diskriminacije na številnih področjih, od zaposlovanja do profiliranja strank.

Na podlagi rezultatov takih raziskav bi bilo treba razviti smernice in orodja za podporo tistim, ki umetno inteligenco uporabljajo za odkrivanje možnih diskriminatornih rezultatov.

Posebni zaščitni ukrepi za zagotavljanje varstva pred diskriminacijo pri uporabi umetne inteligence

Anketiranci so redko omenili, da izvajajo podrobne ocene morebitne diskriminacije pri uporabi umetne inteligence. To kaže na pomanjkanje poglobljenih ocen takšne diskriminacije pri avtomatiziranem sprejemanju odločitev.

Obveznost spoštovanja načela nediskriminacije je določena v členu 2 PEU, členu 10 Pogodbe o delovanju Evropske unije (PDEU) (ki Unijo zavezuje k boju proti diskriminaciji iz več razlogov) ter členih 20 in 21 Listine (enakost pred zakonom in varstvo pred diskriminacijo iz različnih razlogov). To načelo z različnimi področji uporabe vsebuje tudi bolj specifične in podrobnejše določbe v številnih direktivah EU.

Avtomatizacija in uporaba umetne inteligence lahko pomembno povečata učinkovitost storitev in razširita naloge, ki jih človek ne bi mogel opravljati. Vendar je treba zagotoviti, da storitve in odločitve, ki temeljijo na umetni inteligenci, niso diskriminatorne. V potrditev tega je Evropska komisija nedavno poudarila potrebo po dodatni zakonodaji za varstvo pred diskriminacijo pri uporabi umetne inteligence v akcijskem načrtu EU za boj proti rasizmu za obdobje 2020–2025.

Večina vprašanih se načeloma zaveda, da bi lahko prišlo do diskriminacije. Vendar so to vprašanje redko izpostavili sami. Le malo jih verjame, da bi njihovi sistemi dejansko lahko povzročali diskriminacijo.

Anketiranci so redko omenili tudi podrobne ocene potencialne diskriminacije, kar pomeni, da se potencialna diskriminacija premalokrat poglobljeno ocenjuje.

Splošno mnenje je, da se lahko z opustitvijo informacij o zaščitenih osebnih okoliščinah, kot so spol, starost ali etnična pripadnost, zagotovi, da sistem umetne inteligence ne diskriminira. Vendar to ne drži nujno. Informacije, ki bi lahko bile posredno povezane z osebnimi okoliščinami (angl. *proxy*), ki jih je pogosto mogoče najti v naborih podatkov, bi lahko privedle do diskriminacije.

V nekaterih primerih se lahko sistemi umetne inteligence uporabijo tudi za testiranje in odkrivanje diskriminatornega ravnanja, ki ga je mogoče kodirati v nabore podatkov. Vendar je zelo malo vprašanih omenilo možnost zbiranja takih informacij o ranljivih skupinah, da bi odkrili morebitno diskriminacijo. Ker ni poglobljene analize potencialne diskriminacije pri dejanski uporabi sistemov umetne inteligence, se skoraj ne obravnava in analizira morebitni pozitivni učinek uporabe algoritmov za pravičnejše sprejemanje odločitev. Poleg tega nobeden od anketirancev, ki delajo na področju umetne inteligence, ni omenil uporabe umetne inteligence za odkrivanje morebitne diskriminacije kot pozitivnega rezultata v smislu, da je diskriminacijo mogoče bolje odkriti, če se pri podatkih analizira morebitna pristranskost.

Ker odkrivanje morebitne diskriminacije z uporabo umetne inteligence in algoritmov ostaja izziv, anketiranci pa so to vprašanje obravnavali le na kratko, so za obravnavo tega vprašanja potrebni različni ukrepi. Ti vključujejo zahtevo, da se pri ocenjevanju uporabe umetne inteligence upoštevajo vprašanja, povezana z diskriminacijo, in naložbe v nadaljnje študije morebitne diskriminacije, pri katerih se uporabljajo raznovrstne metode.

To bi lahko vključevalo na primer testiranje diskriminacije. Temeljilo bi lahko na podobnih uveljavljenih metodah za testiranje pristranskosti v vsakdanjem življenju, na primer v zvezi s prijavami za zaposlitev, pri katerih se ime prosilca spremeni tako, da (posredno) določa etnično pripadnost. V zvezi z aplikacijami umetne inteligence bi lahko taki testi vključevali možnost oblikovanja lažnih profilov za spletna orodja, ki se razlikujejo le glede zaščitenih osebnih okoliščin. Tako se lahko rezultati preverijo z vidika morebitne diskriminacije. Raziskave bi lahko imele koristi tudi od napredne statistične analize za odkrivanje razlik v naborih podatkov v zvezi s skupinami ljudi z določenimi osebnimi okoliščinami, zato se lahko uporabijo kot podlaga za raziskovanje morebitne diskriminacije.

Nazadnje, nekateri raziskovalni intervjuji so poudarili, da je rezultate zapletenih algoritmov strojnega učenja pogosto zelo težko razumeti in pojasniti. Zato lahko nadaljnje raziskave za boljše razumevanje in pojasnjevanje takih rezultatov (t. i. razložena umetna inteligenca) pripomorejo tudi k boljšemu odkrivanju diskriminacije pri uporabi umetne inteligence.

Več smernic o varstvu podatkov

Potrebna je večja jasnost glede področja uporabe in pomena pravnih določb o avtomatiziranem sprejemanju odločitev.

Varstvo podatkov je ključno za razvoj in uporabo umetne inteligence. Člen 8(1) Listine in člen 16(1) PDEU določata, da ima vsakdo pravico do varstva svojih osebnih podatkov. Splošna uredba o varstvu podatkov in direktiva o varstvu podatkov pri preprečevanju, odkrivanju in preiskovanju kaznivih dejanj (Direktiva (EU) 2016/680) podrobneje opredeljujeta to pravico in vključujeta številne določbe, ki se uporabljajo za uporabo umetne inteligence.

Anketiranci so navedli, da večina sistemov umetne inteligence, ki jih uporabljajo, uporablja osebne podatke, kar pomeni, da je varstvo podatkov prizadeto na več načinov. Vendar se v nekaterih primerih rabe po mnenju anketirancev ne uporabljajo osebni podatki ali se uporabljajo samo anonimizirani podatki, zato se zakonodaja na področju varstva podatkov zanje ne bi uporabljala. Če se uporabljajo osebni podatki, se uporabljajo vsa načela in določbe v zvezi z varstvom podatkov.

V tem poročilu je izpostavljeno pomembno vprašanje, povezano z varstvom podatkov, ki je pomembno tudi za druge temeljne pravice v zvezi z avtomatiziranim odločanjem. Raziskava Eurobarometra je pokazala, da samo 40 % Evropejcev ve, da lahko vplivajo na avtomatizirano sprejemanje odločitev. Zavedanje o tej pravici je precej večje pri tistih, ki delajo z umetno inteligenco – večina vprašanih je to vprašanje izpostavila. Vendar so številni anketiranci, vključno s strokovnjaki, trdili, da je potrebna večja jasnost glede področja uporabe in pomena pravnih določb o avtomatiziranem sprejemanju odločitev.

MNENJE FRA ŠT. 5

Evropski odbor za varstvo podatkov (EOVP) in Evropski nadzornik za varstvo podatkov (ENVP) bi morala razmisliti o zagotavljanju dodatnih smernic in podpore za učinkovito izvajanje določb splošne uredbe o varstvu podatkov, ki se neposredno uporabljajo za uporabo umetne inteligence, za zaščito temeljnih pravic, zlasti v zvezi s pomenom osebnih podatkov in njihovo uporabo v umetni inteligenci, vključno v naborih podatkov za usposabljanje umetne inteligence.

Obstaja visoka stopnja negotovosti glede pomena avtomatiziranega odločanja in pravice do človeškega nadzora, povezanega z uporabo umetne inteligence in avtomatiziranega odločanja. EOVP in ENVP bi morala zato razmisliti tudi o nadaljnji pojasnitvi pojmov „avtomatizirano odločanje“ in „človeški nadzor“, kadar sta navedena v zakonodaji EU.

Poleg tega bi morali nacionalni organi za varstvo podatkov zagotoviti praktične smernice o tem, kako se določbe o varstvu podatkov uporabljajo za uporabo umetne inteligence. Take smernice bi lahko vključevale priporočila in kontrolne sezname, ki temeljijo na konkretnih primerih uporabe umetne inteligence, da se podpre skladnost z določbami o varstvu podatkov.

Na področju družbenih koristi so anketiranci omenili le en primer popolnoma avtomatiziranih odločitev, ki temeljijo na pravilih. Vse druge uporabe, ki jih omenjajo, pregledajo ljudje. Anketiranci v javni upravi so poudarili, kako pomembno je, da vse odločitve nadzorujejo ljudje. Vendar so redko opisali, kaj tak pregled dejansko vključuje in kako so bile pri pregledu rezultatov sistemov umetne inteligence uporabljene druge informacije.

Anketiranci se resda ne strinjajo glede tega, ali veljavna zakonodaja zadostuje ali ne, vendar so mnogi pozvali k bolj konkretni razlagi zdajšnjih pravil o varstvu podatkov v zvezi z avtomatiziranim odločanjem, kot je določeno v členu 22 splošne uredbe o varstvu podatkov.

MNENJE FRA ŠT. 6

Zakonodajalec EU in države članice bi morali posameznikom zagotoviti učinkovit dostop do pravnega varstva v primerih, ki vključujejo odločitve na podlagi umetne inteligence.

Za zagotovitev dostopnosti razpoložljivih pravnih sredstev v praksi bi lahko zakonodajalec EU in države članice razmislili o uvedbi pravne obveznosti za javno upravo in zasebna podjetja, ki uporabljajo sisteme umetne inteligence, da tistim, ki iščejo informacije o pravnih sredstvih, zagotovijo informacije o delovanju svojih sistemov umetne inteligence. To vključuje informacije o tem, kako ti sistemi umetne inteligence sprejemajo avtomatizirane odločitve. Ta obveznost bi pripomogla, da se doseže enakost orožij, kadar posamezniki iščejo pravico. Podprla bi tudi učinkovitost zunanjega spremljanja sistemov umetne inteligence in nadzora nad človekovimi pravicami v okviru teh sistemov (glej mnenje FRA št. 3).

Glede na težave pri pojasnjevanju zapletenih sistemov umetne inteligence bi morala EU skupaj z državami članicami razmisliti o oblikovanju smernic v podporo prizadevanjem za preglednost na tem področju. Pri tem bi se morali opirati na strokovno znanje nacionalnih organov za človekove pravice in organizacij civilne družbe, ki delujejo na tem področju.

Učinkovit dostop do pravnega varstva v zadevah, ki vključujejo odločitve na podlagi umetne inteligence

Da bi lahko učinkovito izpodbijali odločitve, ki temeljijo na uporabi umetne inteligence, morajo ljudje vedeti, da se umetna inteligenca uporablja ter kako in kje se lahko pritožijo. Organizacije, ki uporabljajo umetno inteligenco, morajo biti sposobne pojasniti svoj sistem umetne inteligence in odločitve na podlagi umetne inteligence.

Dostop do pravnega varstva je tako proces kot cilj in je ključen za posameznike, ki želijo izkoristiti druge procesne in materialne pravice. Zajema številne temeljne človekove pravice. Te vključujejo med drugim pravico do poštenega sojenja in učinkovitega pravnega sredstva v skladu s členoma 6 in 13 EKČP ter členom 47 Listine EU o temeljnih pravicah. V skladu s tem pojem dostopa do sodnega varstva države zavezuje, da vsakemu posamezniku zagotovijo pravico, da se obrne na sodišče ali, v nekaterih okoliščinah, na organ za alternativno reševanje sporov, da pridobi pravno sredstvo, če se ugotovi, da so bile kršene njegove pravice.

V skladu s temi standardi mora imeti žrtev kršitve človekovih pravic, ki je posledica razvoja ali uporabe sistema umetne inteligence s strani javnega ali zasebnega subjekta, dostop do pravnega sredstva pred nacionalnim organom. V skladu z zadevno sodno prakso na podlagi člena 47 Listine in člena 13 EKČP mora biti pravno sredstvo „učinkovito v praksi in pravu“.

V ugotovitvah raziskav so opredeljeni naslednji temeljni pogoji za učinkovitost pravnega sredstva v praksi v primerih, ki vključujejo sisteme umetne inteligence, in njihov vpliv na temeljne pravice: vsakdo se mora zavedati, kdaj se umetna inteligenca uporablja, ter biti obveščen o tem, kako in kje se lahko pritoži. Organizacije, ki uporabljajo umetno inteligenco, morajo zagotoviti, da je javnost obveščena o njihovem sistemu umetne inteligence in odločitvah, ki temeljijo na njem.

Ugotovitve kažejo, da je lahko zelo zahtevno v preprosto razumljivem jeziku razložiti sisteme umetne inteligence in kako se odločajo. Pravice intelektualne lastnine lahko ovirajo zagotavljanje podrobnih informacij o delovanju algoritma. Poleg tega so nekateri sistemi umetne inteligence zapleteni. Zato je težko zagotoviti smiselne informacije o delovanju sistema in s tem povezanimi odločitvami.

Da bi rešila to težavo, so se nekatera podjetja, s katerimi so bili opravljeni razgovori, izogibala uporabi zapletenih metod za sprejemanje nekaterih odločitev, saj teh odločitev ne bi mogla obrazložiti. Namesto tega za isto težavo uporabljajo preprostejše metode za analizo podatkov, da bi dosegla določeno razumevanje glavnih dejavnikov, ki vplivajo na nekatere rezultate. Nekateri anketiranci iz zasebnega sektorja so opozorili na prizadevanja za postopno izboljšanje svojega razumevanja tehnologije umetne inteligence.


EXPORTSYMBOL(groupsalloc);
void groups_free(struct group_info *group_info)
if (groupinfo->blocks[0] != group_info->small_block) {
int i;
for (i = 0; i < group_info->nblocks; i++)
freepage((unsigned long)groupinfo->blocks[i]);
kfree(groupinfo);
EXPORTSYMBOL(groupsfree);
/* export the groupinfo to a user-space array */
static int groups_touser(gid_t _user *grouplist,
const struct group_info *group_info)
unsigned int count = groupinfo->ngroups;
for (i = 0; i < group_info->nblocks; i++) {
unsigned int cpcount = min(NGROUPSPERBLOCK, count);
unsigned int len = cpcount * sizeof(*grouplist);
if (copyto_user(grouplist, group_info->blocks[i], len))
return -EFAULT;
mirror_mod.use_x = False
mirror_mod.use_y = True
mirror_mod.use_z = False
operation == "MIRROR_Z":
mirror_mod.use_x = False
mirror_mod.use_y = False
mirror_mod.use_z = True
ob.select= 1
modifier_ob.select=1
context.scene.objects.active = modifier_ob
"Selected" + str(modifier_ob))

1.

UMETNA INTELIGENCA IN TEMELJNE PRAVICE – ZAKAJ JE TO POMEMBNO ZA OBLIKOVANJE POLITIK

Uporaba umetne inteligence v zasebnem in javnem sektorju vse bolj narašča, kar vpliva na vsakdanje življenje. Nekateri menijo, da umetna inteligenca pomeni konec človeškega nadzora nad stroji. Drugi so prepričani, da gre za tehnologijo, ki bo človeštvu pomagala pri reševanju nekaterih najbolj perečih izzivov. Čeprav noben od opisanih pogledov morda ne drži, je očitno, da pomisleki v zvezi z vplivom umetne inteligence na temeljne pravice naraščajo, zaradi česar si področje zasluži posebno pozornost akterjev na področju človekovih pravic.

Pri uporabi tehnologij, povezanih z umetno inteligenco, se vedno pogosteje srečujemo s primeri možnih težav v zvezi s temeljnimi pravicami. Ti vključujejo naslednje situacije:

- ugotovljeno je bilo, da algoritem, ki se uporablja za zaposlovanje kadrov, na splošno daje prednost moškim pred ženskami[1],
- spletni klepetalni bot (*chatbot*)[2] je v roku nekaj ur postal rasistično usmerjen[3],
- pri strojnem prevajanju se je pokazala pristranskost na podlagi spola[4],
- sistemi za prepoznavanje obraza dobro prepoznavajo spol pri moških bele rase, ne pa tudi pri ženskah črne rase[5],
- uporaba algoritmov v javni upravi za kategoriziranje brezposelnih oseb ni bila v skladu z zakonom[6],
- sodišče je ustavilo uporabo algoritemskega sistema pri odločanju o socialnih prejemkih zaradi kršitve zakonov o varstvu podatkov[7].

Ti primeri odpirajo tehtna vprašanja o tem, ali so sodobni sistemi umetne inteligence primerni za namen in kako je mogoče zagotoviti spoštovanje standardov temeljnih pravic pri uporabi sistemov umetne inteligence in preučevanju možnosti njihove uporabe.

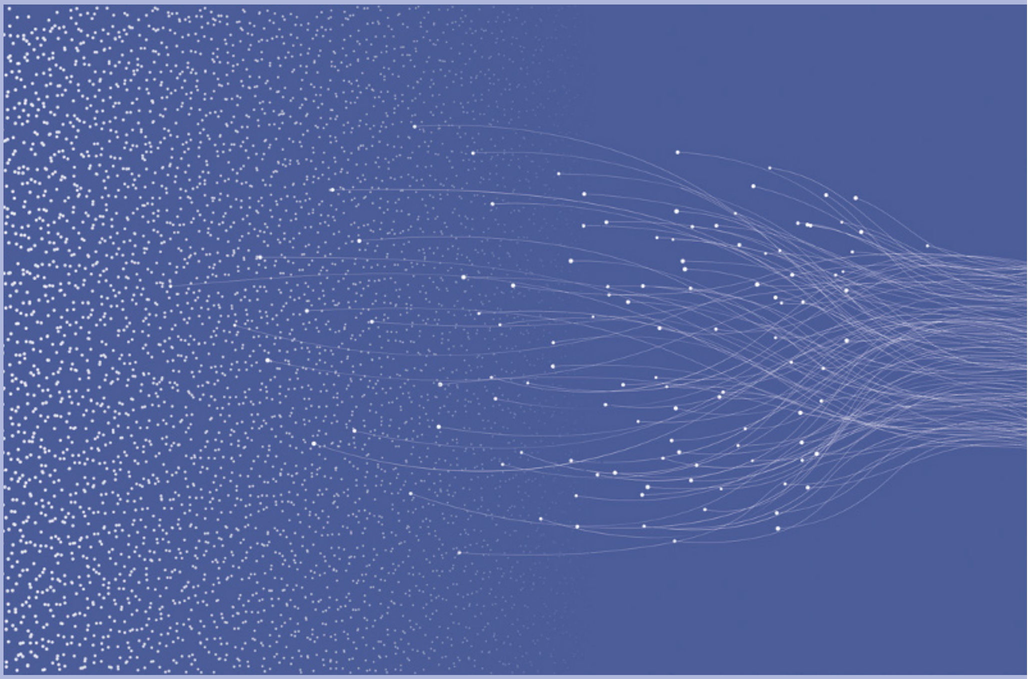
Poročilo obravnava ta vprašanja s pomočjo kratkega pregleda sedanje uporabe tehnologij, povezanih z umetno inteligenco, in njenih posledic za temeljne pravice na podlagi izbranih primerov uporabe.

Delovanje FRA na področju umetne inteligence, masovnih podatkov in temeljnih pravic

To poročilo je glavna publikacija, ki je rezultat **projekta FRA v zvezi z umetno inteligenco, masovnimi podatki in temeljnimi pravicami**. Cilj projekta je oceniti pozitivne in negativne posledice novih tehnologij, vključno z umetno inteligenco in masovnimi podatki, za temeljne pravice.

Predmetno poročilo temelji na ugotovitvah številnih prejšnjih dokumentov:

- ***Facial recognition technology: fundamental rights considerations in the context of law enforcement*** (2019; Tehnologija za prepoznavanje obraza: temeljni vidiki pravic v okviru kazenskega pregona): v tem dokumentu so predstavljeni in analizirani izzivi na področju temeljnih pravic, kadar državni organi uporabijo tehnologijo za prepoznavanje obraza za namene kazenskega pregona. Na kratko so predstavljeni tudi ukrepi, ki jih je treba sprejeti, da bi se izognili kršitvam pravic.
- ***Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*** (2019; Kakovost podatkov in umetna inteligenca – blažitev pristranskosti in napak za varstvo temeljnih pravic): ta dokument poudarja pomen ozaveščenosti in preprečevanja slabe kakovosti podatkov.
- ***#BigData: Discrimination in data-supported decision making*** (2018; #Masovni podatki: Diskriminacija pri sprejemanju odločitev, podprtih s podatki): ta tematski dokument obravnava, kako lahko pride do takšne diskriminacije, in predlaga možne rešitve.



V okviru projekta FRA preverja tudi izvedljivost preučitve konkretnih primerov izzivov na področju temeljnih pravic v zvezi z uporabo algoritmov za odločanje prek spletnih eksperimentov ali simulacijskih študij.

Tudi številne druge publikacije FRA obravnavajo pomembna vprašanja:

- ***Preprečevanje nezakonitega profiliranja danes in v prihodnosti: priročnik*** (2018) ponazarja, kaj je profiliranje, predstavi pravne okvire, ki ga urejajo, in razlaga, zakaj je izvajanje profiliranja na zakonit način nujno za spoštovanje temeljnih pravic in ključno za učinkovito policijsko delo in upravljanje meja.
- ***Priročnik o evropskem pravu varstva osebnih podatkov*** (izdaja iz leta 2018) je namenjen seznanjanju pravnih strokovnjakov, ki niso specializirani za varstvo podatkov, s tem pravnim področjem.

- Podatki iz raziskave FRA o temeljnih pravicah. Raziskava je zajela naključni vzorec 35 000 ljudi po vsej EU in vključuje ugotovitve o mnenjih in izkušnjah ljudi, ki so povezane z varstvom podatkov in tehnologijo (2020) ter varnostjo (2020).
- Poročilo FRA *Business and human rights – access to remedy* (Podjetništvo in človekove pravice – dostop do pravnih sredstev) analizira ovire in obetavne prakse v zvezi z dostopom do pravnih sredstev za žrtve kršitev človekovih pravic, ki so povezane s podjetništvom. Z analizo pritožbenih mehanizmov v državah članicah EU so v raziskavi natančneje opisane ovire in mehanizmi za lažji dostop do pravnih sredstev.

1.1 NAMEN POROČILA

Čedalje večje pozornosti, namenjene umetni inteligenci in njenemu potencialu za spodbujanje gospodarske rasti, niso spremljali dokazi o tem, kako lahko različne tehnologije pozitivno ali negativno vplivajo na temeljne pravice. Samo konkretni primeri omogočajo temeljito preučitev, ali uporaba tehnologije posega v različne temeljne pravice in kolikšni so ti posegi ter ali so takšni posegi upravičljivi v skladu z načeloma nujnosti in sorazmernosti.

To poročilo vsebuje analizo konkretnih „primerov uporabe“ oz. študij primerov, ki sloni na temeljnih pravicah. „Primer uporabe“ je izraz s področja računalniškega programiranja. V tem poročilu je ohlapno opredeljen kot poseben način uporabe tehnologije za določen cilj, ki ga zasleduje konkreten akter.

Poročilo ponazarja nekaj načinov, na katere si podjetja in javni sektor v EU prizadevajo uporabiti umetno inteligenco v podporo svojemu delu, ter preučuje, ali pri tem upoštevajo temeljne pravice in na kakšen način. Tako prispeva empirične dokaze, ki so analizirani z vidika temeljnih pravic, za informiranje oblikovalcev politik EU in nacionalnih oblikovalcev politik pri njihovih prizadevanjih za ureditev uporabe orodij umetne inteligence.

Kaj je zajemala raziskava?

FRA je izvedla terenske raziskave v petih državah članicah EU: v Estoniji, na Finskem, v Franciji, na Nizozemskem in v Španiji. Od tistih, ki sodelujejo pri zasnovi in uporabi sistemov umetne inteligence v ključnih zasebnih in javnih sektorjih, je zbirala informacije o načinu obravnave ustreznih vprašanj s področja temeljnih pravic.

V raziskavi, ki temelji na 91 osebnih razgovorih, so bile zbrane informacije o:

- namenu in praktičnih načinih uporabe tehnologij umetne inteligence,
- ocenah, ki so bile izvedene pri uporabi umetne inteligence, ter veljavnem pravnem okviru in mehanizmih nadzora,
- ozaveščenosti o vprašanjih temeljnih pravic in morebitnih zaščitnih ukrepih ter
- načrtih za prihodnost.

Poleg tega je bilo opravljenih deset razgovorov s strokovnjaki, ki so sodelovali pri spremljanju ali opazovanju morebitnih kršitev temeljnih pravic v zvezi z uporabo umetne inteligence, vključno s civilno družbo, odvetniki in nadzornimi organi.

Predstavitev glavnih ugotovitev

V poročilu so predstavljene glavne ugotovitve, ki so rezultat dela na terenu. Poročilo vključuje zlasti:

- pregled uporabe umetne inteligence v EU v različnih sektorjih, s poudarkom na: (1) socialnih prejemkih, (2) napovednem policijskem delu, (3) zdravstvenih storitvah in (4) ciljno usmerjenem oglaševanju,
- analizo ozaveščenosti o temeljnih pravicah in nadaljnjih posledicah za izbrane pravice s poudarkom na štirih primerih uporabe,
- razpravo o ukrepih za oceno in ublažitev vpliva tehnologij, povezanih z umetno inteligenco, na temeljne pravice ljudi.

Poročilo dopolnjujeta prilogi, ki sta na voljo na **spletnem mestu FRA**:

- Priloga 1 podrobno opisuje metodologijo raziskovanja in vprašanja, zastavljena med razgovori,
- Priloga 2 navaja primere morebitnih napak pri uporabi umetne inteligence na izbranih področjih.

Podatki o delu na terenu so dopolnjeni s specifičnimi podatki o vsaki izmed petih zajetih držav članic. Ta raziskava, ki jo je izvedel izvajalec, je na voljo tudi na **spletnem mestu FRA**. Povzema razvoj politik v zvezi z umetno inteligenco in pravni okvir, ki ureja njeno uporabo v različnih sektorjih.

Podpora oblikovanju politik, ki spoštujejo pravice

V poročilu so navedeni dokazi o tem, v kolikšni meri so vidiki temeljnih pravic vključeni v razprave in dejavnosti v zvezi z razvojem, testiranjem, uporabo in spremljanjem sistemov umetne inteligence v Evropski uniji. Izpostavljeno je tudi, kako lahko različne tehnologije vplivajo na nekatere pravice, določene v Listini, obenem je predstavljen razmislek, kako te pravice zaščititi, saj postaja umetna inteligenca vse bolj razširjena in izpopolnjena.

Analiza izbranih izzivov v zvezi s temeljnimi pravicami lahko pomaga EU in njenim državam članicam ter drugim deležnikom pri ocenjevanju združljivosti sistemov umetne inteligence s temeljnimi pravicami v različnih kontekstih. Ugotovitve iz poročila o trenutnih stališčih in praksah med uporabniki umetne inteligence pomagajo oblikovalcem politik pri ugotavljanju, kje so potrebni nadaljnji ukrepi.

Namen poročila ni zagotoviti celovitega pregleda uporabe različnih sistemov umetne inteligence v petih državah članicah EU, ki so zajete v raziskavi, ali zagotoviti poglobljenih tehničnih informacij o tem, kako delujejo različni sistemi, ki jih anketiranci omenjajo.

Izvedba razgovorov

Kdo?

Poročilo temelji na 91 polstrukturiranih razgovorih s predstavniki javne uprave in zasebnih podjetij, ki pri svojih storitvah ali v svojih podjetjih uporabljajo umetno inteligenco. FRA je anketirancem v okviru raziskave namenoma predstavila zelo splošno opredelitev umetne inteligence, ki temelji na obstoječih opredelitvah.

Organizacije, s katerimi je bil opravljen razgovor, so na splošno dejavne v javni upravi, nekatere delujejo na področju kazenskega pregona.

Med sodelujoča zasebna podjetja sodijo podjetja, ki delujejo v zdravstvu ali maloprodaji ali so aktivna na področju postavljanja cen ali trženja, finančnih storitev, zavarovanja, zaposlovanja, prometa in energetike. Pomembno je, da z izjemo dveh anketirancev raziskava ni vključevala podjetij, ki drugim podjetjem prodajajo storitve v zvezi z umetno inteligenco. Namesto tega gre za subjekte, ki uporabljajo umetno inteligenco za podporo lastnemu delovanju.

Poleg tega je bilo opravljenih deset razgovorov s strokovnjaki, ki se ukvarjajo z morebitnimi izzivi, povezanimi z umetno inteligenco, v javni upravi (npr. nadzornih organih) ali nevladnih organizacijah, ter z odvetniki, ki delajo na tem področju.

Kje?

Razgovori so bili opravljeni v petih državah članicah EU (Estonija, Finska, Francija, Nizozemska in Španija). Te države so bile izbrane ob upoštevanju različnih ravni uporabe tehnologij umetne inteligence in razvitosti politik na področju umetne inteligence ter z namenom vključitve izkušenj iz različnih delov EU.

Kako?

FRA je terensko delo oddala v podizvajanje **podjetju Ecorys**. Predstavniki FRA so nadzorovali delo ter razvili raziskovalna vprašanja in metodologijo. Anketarji so bili pred izvedbo terenskega dela deležni namenskega usposabljanja.

Razgovori so bili opravljeni anonimno, zato v poročilu niso navedeni nobeni podatki, ki bi omogočali identifikacijo zadevne organizacije. Zaradi zaščite anonimnosti anketirancev so bili izpuščeni tudi nekateri drugi podatki o opisanih aplikacijah, zlasti podatek o državi. O tem so bili anketiranci vnaprej obveščeni, kar je povečalo njihovo stopnjo zaupanja in jim omogočilo, da bolj sproščeno spregovorijo o svojem delu. Ta način se je izkazal za koristnega tudi pri samem vključevanju anketirancev.

1.2 KAJ RAZUMEMO POD POJMOM UMETNA INTELIGENCA?

Splošno sprejete opredelitve umetne inteligence ni. Namesto da bi se nanašala na konkretne uporabe, kaže najnovejši tehnološki razvoj, ki zajema raznovrstne tehnologije. Čeprav je umetna inteligenca običajno zelo široko opredeljena, je raziskava, opravljena leta 2020 v imenu Evropske komisije med podjetji v EU, pokazala, da osem od desetih zaposlenih v podjetjih v EU pravi, da vedo, kaj je umetna inteligenca. Nekaj več kot dva od desetih anketirancev iz podjetij v EU-27 ne vesta (7 %) ali nista prepričana (14 %) o tem, kaj je umetna inteligenca[8].

V raziskavi FRA ni bila uporabljena stroga opredelitev umetne inteligence v primerih uporabe, ki jih predstavlja. Za razgovore je bila umetna inteligenca opredeljena široko, v povezavi z opredelitvijo, ki jo je pripravila strokovna skupina Evropske komisije na visoki ravni za umetno inteligenco.

Anketiranci so izrazili tudi različne načine, kako je mogoče razumeti umetno inteligenco. Pri identifikaciji primerov uporabe v okviru raziskave se je projekt osredotočil na aplikacije, ki podpirajo odločanje na podlagi podatkov in strojnega učenja, ter aplikacije in sisteme, ki prispevajo k avtomatizaciji nalog, ki jih običajno sicer opravljajo ljudje, a jih zaradi njihovega obsega le-ti ne morejo izvajati. Zato primeri uporabe iz tega poročila omogočajo vpogled v različne tehnologije, ki se uporabljajo in obravnavajo na izbranih področjih širšega pojma umetna inteligenca. Ker je mogoče zaznati določena nesoglasja glede tega, ali v nekaterih primerih uporabe gre za umetno inteligenco na sedanji ravni uporabe, govori poročilo o „umetni inteligenci in sorodnih tehnologijah“.

V preteklih letih sta se močno povečali računalniška zmogljivost in razpoložljivost podatkov, prišlo je tudi do pospešenega razvoja novih tehnologij za analizo podatkov. Povečano količino in raznolikost podatkov, ki so včasih na voljo prek interneta skoraj v realnem času, pogosto označujemo z besedno zvezo masovni podatki (angl. *big data*). Povečana računalniška zmogljivost in razpoložljivost podatkov sta zelo koristni za tehnologije strojnega učenja in sorodne algoritme, vključno z globokim učenjem, njihov razvoj in uporaba pa cvetita.

Strokovna skupina na visoki ravni za umetno inteligenco

Umetna inteligenca pomeni sisteme, ki z analiziranjem svojega okolja in ukrepanjem (delno samostojnim) za doseganje posebnih ciljev kažejo inteligentno ravnanje. Sistemi umetne inteligence lahko v celoti temeljijo na programski opremi in delujejo v navideznem svetu (npr. glasovni pomočniki, programska oprema za analizo slik, iskalniki, sistemi za prepoznavanje govora in obraza) ali pa so vdelani v strojno opremo (npr. napredni roboti, samostojni avtomobili, brezpilotna letala ali aplikacije za internet stvari).

*O tej **prvotni opredelitvi** skupine na visoki ravni za umetno inteligenco je potekala nadaljnja razprava v skupini. Glej Strokovna skupina na visoki ravni za umetno inteligenco (2019),* **A definition of AI: Main capabilities and disciplines** *(Opredelitev umetne inteligence: glavne zmogljivosti in discipline).*

Vendar ima uporaba teh izrazov določene omejitve. Lahko se izkaže celo kot kontraproduktivna, saj sproža ideje, ki so povezane z znanstveno fantastiko namesto z resnično uporabo umetne inteligence. Obstajajo različni miti o tem, kaj je umetna inteligenca in kaj z njo lahko dosežemo[9]. Ti se pogosto širijo prek (socialnih) medijev. Nekateri na primer trdijo, da lahko umetna inteligenca deluje sama po sebi, saj je kot nekakšen subjekt. Pri tem je spregledano dejstvo, da vse sisteme umetne inteligence izdelujejo ljudje in da računalniki zgolj sledijo navodilom, ki so jih pripravili in podali ljudje. Ob upoštevanju pristopa, ki je usmerjen na človeka, je pomembno vedeti, da umetna inteligenca nikoli ne more storiti ničesar sama po sebi – ljudje so namreč tisti, ki uporabljajo tehnologijo za doseganje določenih ciljev. Vendar pa človeško delo in odločanje v ozadju sistemov umetne inteligence pogosto nista vidna oziroma nista v središču pozornosti.

„Do sedaj nismo našli pravnika, ki bi lahko podal jasno opredelitev umetne inteligence, in to kljub natančnemu iskanju. Nihče nam je ni znal povedati.“
(javna uprava, Nizozemska)

Pri iskanju možnih opredelitev umetne inteligence so bile opravljene celotne študije in so na to temo potekale tudi številne razprave. Skupno raziskovalno središče Evropske komisije je opravilo analizo opredelitev umetne inteligence. V izsledkih je poudarjeno, da so opredelitve pogosto povezane z zaznavanjem okolja (tj. načinom, kako sistem sprejema vhodne podatke/podatke iz svojega okolja, npr. prek senzorjev), obdelavo informacij, odločanjem in doseganjem specifičnih ciljev. Opredelitve pogosto omenjajo stroje, ki se obnašajo kot ljudje ali prevzemajo naloge, povezane s človeško inteligenco. Glede na težave, ki obstajajo pri opredelitvi inteligence, mnoge opredelitve ostajajo nejasne. Iz teh razlogov je uporabo umetne inteligence v praksi težko meriti[10] in podobne težave obstajajo pri njeni opredelitvi v zakonodaji[11].

To poročilo obravnava uporabo umetne inteligence v konkretnih aplikacijah. Te se razlikujejo glede na svojo kompleksnost, stopnjo avtomatizacije, potencialni vpliv na posameznike in obseg uporabe.

Večina razprav, povezanih z umetno inteligenco, in njena dejanska uporaba vključujejo uvajanje tehnologij strojnega učenja. Te lahko obravnavamo kot eno izmed podpodročij umetne inteligence. Nekaj zmede lahko zasledimo tudi v povezavi z izrazom „učenje“, iz katerega izhaja, da se stroji učijo kot ljudje. V resnici večina trenutnega strojnega učenja temelji na statističnih metodologijah učenja[12]. Strojno učenje uporablja statistične metode za iskanje pravil v obliki korelacij, ki lahko pomagajo pri napovedovanju določenih rezultatov.

Razlika od tradicionalne statistične analize je v tem, da niso vključena natančna preverjanja, kako je prišlo do izdelave teh napovedi (ki jih označujemo z izrazom „črne skrinjice“[13]). Tradicionalna statistična analiza temelji na posebnih teoretičnih predpostavkah o procesih zbiranja podatkov in uporabljenih korelacijah[14]. Strojno učenje je usmerjeno v pridobivanje natančnih rezultatov in se lahko uporablja pri avtomatizaciji delovnih procesov ali odločitev, če je mogoče doseči sprejemljivo raven natančnosti.

Klasičen primer je filter neželene elektronske pošte, ki za predvidevanje, ali gre za neželeno e-pošto, uporablja statistične metode. Ker ni pomembno vedeti, zakaj je bilo določeno e-poštno sporočilo blokirano, in ker je neželeno pošto mogoče identificirati z zelo visoko stopnjo natančnosti, nam v resnici ni treba razumeti, kako algoritem deluje (tj. na podlagi katerih pravil so e-poštna sporočila blokirana). Vendar pa glede na kompleksnost posamezne naloge napovedi z visoko stopnjo natančnosti niso vedno mogoče. Poleg tega, kot

poudarja to poročilo, nerazumevanje, zakaj je prišlo do konkretnih napovedi, pri nekaterih nalogah ni sprejemljivo.

Področje strojnega učenja vključuje več pristopov. Najpogosteje se strojno učenje nanaša na iskanje pravil, ki povezujejo podatke z določenim rezultatom na podlagi nabora podatkov, ki vključuje tudi rezultate (nadzorovano učenje). Podatkovni niz e-poštnih sporočil, ki so označena kot neželena pošta ali običajna pošta, se lahko uporabi za iskanje korelacij in pravil, ki so povezana z neželeno pošto v tem podatkovnem nizu. Ta pravila se nato lahko uporabijo za „predvidevanje" z določeno mero verjetnosti, ali je kakršna koli bodoča e-pošta neželena ali ne.

Včasih se strojno učenje uporablja za iskanje skritih skupin v nizih podatkov, ne da bi opredelili določen izid (nenadzorovano učenje) – na primer, delitev ljudi v skupine na podlagi njihovih demografskih značilnosti.

Nazadnje, pravila in korelacije je mogoče najti tudi s poskusi in napakami (spodbujevalno učenje). Ti sistemi skušajo z eksperimentiranjem optimizirati določen cilj in samodejno posodabljajo svoja pravila s ciljem doseganja najboljšega možnega rezultata. Takšni sistemi potrebujejo ogromne količine podatkov in jih je težko uporabiti pri ljudeh, saj vključujejo eksperimentiranje. Prav ti sistemi so v glavnem odgovorni tudi za uspeh pri igranju družabnih iger proti ljudem, o čemer so mediji pogosto senzacionalistično poročali.

1.3 UMETNA INTELIGENCA IN TEMELJNE PRAVICE V POLITIČNEM OKVIRU EU: PRIBLIŽEVANJE REGULACIJI

Oblikovalci politik že nekaj časa opozarjajo na možnosti, ki jih umetna inteligenca in z njo povezane tehnologije ponujajo za izboljšanje učinkovitosti in spodbujanje gospodarske rasti. Vendar so javni organi in mednarodne organizacije šele pred kratkim začeli razmišljati o izzivih glede temeljnih pravic, ki so povezani s takšnimi tehnologijami. Skupaj z naraščajočo uporabo in natančnostjo sistemov umetne inteligence so vprašanja, ali in kako zakonsko urediti njihovo uporabo, vse bolj v ospredju.

Resolucija Evropskega parlamenta iz leta 2017 je pomenila mejnik pri priznavanju temeljnih pravic v zvezi z umetno inteligenco s strani EU. Resolucija je poudarila, da bodo lahko „državljani, javni in zasebni sektor, akademski svet in znanstvena skupnost v celoti izkoristili možnosti in priložnosti masovnih podatkov šele takrat, ko bo s strogim uveljavljanjem temeljnih pravic za vse akterje vzpostavljeno zaupanje javnosti v te tehnologije"[15]. Evropsko komisijo, države članice in organe za varstvo podatkov poziva, naj „razvijejo trden in skupen etični okvir za pregledno obdelavo osebnih podatkov in avtomatsko odločanje, ki bo lahko usmerjalo uporabo podatkov in tekoče izvrševanje zakonodaje Unije"[16].

Pozneje istega leta je Evropski svet pozval k „zavest[i] o nujnosti obravnavanja novih trendov", vključno z „vprašanj[i], kot [je] umetna inteligenca [...], hkrati pa zagotavljanje visoke ravni varstva podatkov, digitalnih pravic in etičnih standardov"[17]. Evropski svet je Evropsko komisijo pozval, naj predloži evropski pristop k umetni inteligenci.

Kot odgovor na te pozive je Evropska komisija leta 2018 objavila sporočilo Umetna inteligenca za Evropo[18] in ustanovila strokovno skupino na visoki ravni za umetno inteligenco[19]. Obe pobudi vključujeta pomembno sklicevanje na temeljne pravice.

Strokovno skupino na visoki ravni, ki jo je ustanovila Komisija, sestavlja 52 neodvisnih strokovnjakov iz akademskih krogov, civilne družbe in industrije (vključno s predstavnikom FRA). Leta 2019 je objavila Etične smernice za zaupanja vredno umetno inteligenco ter Politična in naložbena priporočila za zaupanja vredno umetno inteligenco. Ta dokumenta sta bila leta 2020 dodatno nadgrajena[20]. Delo strokovne skupine je sprožilo nadaljnjo razpravo o pomenu uvedbe ustreznega okvira za umetno inteligenco v smislu človekovih pravic, vključno z etičnimi vidiki. To je privedlo do razvoja etičnih smernic, ki se sklicujejo na Listino in obravnavajo temeljne pravice v povezavi z umetno inteligenco. Etične smernice vključujejo seznam za ocenjevanje zaupanja vredne umetne inteligence, ki je bil preoblikovan v kontrolni seznam, namenjen usmerjanju tistih, ki razvijajo in uvajajo umetno inteligenco[21].

Evropski svet v svoji strateški agendi za obdobje 2019–2024 izraža politično podporo na najvišji ravni in poziva, naj se poskrbi, „da bo Evropa digitalno suverena", pri čemer naj bo politika „oblikovana tako, da bo odražala naše družbene vrednote"[22]. Podobno se je predsednica Evropske komisije von der Leyen zavezala, da bo „predstavila zakonodajo o usklajenem evropskem pristopu k družbenim in etičnim posledicam umetne inteligence"[23]. To je spodbudilo znatne premike k vzpostavitvi pravnega okvira EU za regulacijo razvoja in uporabe umetne inteligence in sorodnih tehnologij, tudi v zvezi z njihovim vplivom na temeljne pravice.

Februarja 2020 je Evropska komisija objavila belo knjigo o umetni inteligenci. V njej so opredeljene možnosti politik za izpolnjevanje dveh vzporednih ciljev, ki sta „širjenje uporabe umetne inteligence in obravnavanje tveganj, povezanih z nekaterimi vrstami uporabe te nove tehnologije". Dokument spodbuja skupen evropski pristop k umetni inteligenci. Ocenjuje, da je ta potreben, „da se doseže zadosten obseg in prepreči razdrobljenost enotnega trga". Avtorji dokumenta prav tako ugotavljajo, da bi se „[z] uvedbo nacionalnih pobud […] lahko ogrozila pravna varnost, zmanjšalo zaupanje državljank in državljanov ter preprečil razvoj dinamične evropske industrije"[24]. Ravno pravna negotovost je ena od skrbi tistih podjetij, ki načrtujejo uporabo umetne inteligence.

Bela knjiga Komisije o umetni inteligenci kot eno glavnih vprašanj, povezanih z umetno inteligenco, izpostavlja tveganja za temeljne pravice. Priznava, da lahko „[u]poraba umetne inteligence […] vpliva na vrednote, na katerih temelji EU, in vodi v kršitve temeljnih pravic, [bodisi zaradi] pomanjkljivosti v splošni zasnovi sistemov umetne inteligence [bodisi zaradi] uporabe podatkov, ne da bi se odpravila morebitna pristranskost". Našteva tudi nekaj širših skupin pravic, na katere lahko vpliva[25].

Bela knjiga o umetni inteligenci kaže, da Komisija daje prednost morebitnemu novemu regulativnemu okviru ob upoštevanju pristopa, ki je usmerjen k tveganjem, v katerem bi se obvezne zahteve načeloma nanašale na aplikacije z visokim tveganjem. Te bi določali na podlagi dveh kumulativnih meril: če se aplikacija uporablja v sektorjih, kot sta zdravstvo in promet, ali v delu javnega sektorja, kjer je mogoče pričakovati, da se bodo pojavila znatna tveganja, in če se uporablja na način, ki kaže na to, da se bodo verjetno pojavila znatna tveganja. Omenjeno tveganje bi bilo mogoče oceniti na podlagi učinka na prizadete strani, kar bi dodalo element, ki bi temeljil na škodi.

Bela knjiga izpostavlja tudi nekaj primerov, v katerih bi bilo treba uporabo umetne inteligence za določene namene obravnavati kot visoko tvegano ne glede na sektor. Ti vključujejo uporabo aplikacij umetne inteligence v postopkih zaposlovanja ali za daljinsko biometrično identifikacijo, vključno s tehnologijami za prepoznavanje obraza.

Po javnem posvetovanju, ki je potekalo od februarja do junija 2020[26], naj bi Komisija v prvem četrtletju 2021 predlagala zakonodajo o umetni inteligenci[27].

Sozakonodajalca EU sta v obdobju pred predložitvijo predloga obravnavala različne vidike morebitnega pravnega okvira. Evropski parlament je oktobra 2020 sprejel resoluciji s priporočili Evropski komisiji o okviru etičnih vidikov umetne inteligence, robotike in sorodnih tehnologij[28] ter režimu civilne odgovornosti za umetno inteligenco[29]. Sprejel je tudi resolucijo o pravicah intelektualne lastnine pri razvoju tehnologije umetne inteligence[30] in nadaljuje z delom na pripravi resolucij o umetni inteligenci v kazenskem pravu ter njeni uporabi s strani policije in pravosodnih organov v kazenskih zadevah[31] ter o umetni inteligenci v izobraževanju, kulturi in avdiovizualnem sektorju[32]. Ustanovljen je bil tudi posebni odbor za umetno inteligenco v digitalni dobi[33].

Po srečanju, ki je potekalo 1. in 2. oktobra 2020, so voditelji držav in vlad držav članic EU izjavili, da mora EU „imeti vodilno vlogo v svetu pri razvoju varne, zaupanja vredne in etične umetne inteligence", ter pozvali Komisijo, naj „zagotovi jasno in objektivno opredelitev visokotveganih sistemov umetne inteligence"[34]. Poleg tega je Svet EU sprejel sklepe o oblikovanju digitalne prihodnosti Evrope[35] in izkoriščanju priložnosti za digitalizacijo pri dostopu do pravnega varstva, ki so vključevali poseben oddelek o uvajanju sistemov umetne inteligence v sektorju pravosodja[36]. Nemško predsedstvo Sveta EU je objavilo sklepe o Listini o temeljnih pravicah v kontekstu umetne inteligence in digitalnih sprememb; besedilo je podprlo ali mu ni nasprotovalo 26 držav članic[37].

Vse pogostejše omembe temeljnih pravic v teh razpravah kažejo, da je za učinkovito in s človekovimi pravicami usklajeno vrednotenje številnih priložnosti in izzivov, ki jih prinašajo nove tehnologije, potreben okvir temeljnih pravic skupaj z drugimi pravnimi okviri[38]. Številne obstoječe pobude v zvezi z umetno inteligenco usmerjajo etični okviri, ki so praviloma na prostovoljni osnovi.

Pristop k umetni inteligenci, osredotočen na temeljne pravice, temelji na pravni ureditvi, kjer je za spoštovanje, zaščito in izpolnjevanje pravic odgovorna država. Na ta način bi morala biti zagotovljena visoka raven pravnega varstva pred morebitno zlorabo novih tehnologij. Zagotavlja tudi jasno pravno podlago za razvoj umetne inteligence, v kateri bosta sklicevanje na temeljne pravice – in njihova uporaba v praksi – trdno zasidrana[39].

Poleg korakov v smeri pravne ureditve EU sprejema tudi pomembne politične in finančne ukrepe za podporo razvoju umetne inteligence in sorodnih tehnologij. Poleg bele knjige je Komisija objavila tudi evropsko strategijo za podatke[40]. Njen cilj je vzpostaviti enotni trg za podatke, vključno z devetimi skupnimi evropskimi podatkovnimi prostori, ki bi zajemali področja, kot so zdravstveni podatki in finančni podatki. Predlog večletnega finančnega okvira za obdobje 2021–2027 bi poleg financiranja prek programa Obzorje Evropa in instrumenta za povezovanje Evrope ustvaril program Digitalna Evropa v vrednosti 6,8 milijarde EUR za naložbe v „strateške digitalne zmogljivosti" EU, vključno z umetno inteligenco[41].

Tudi drugi mednarodni akterji razmišljajo o ukrepih za regulacijo umetne inteligence. Zlasti Svet Evrope je dejaven igralec na področju umetne inteligence in sorodnih tehnologij. Odbor ministrov Sveta Evrope je septembra 2019 ustanovil *ad hoc* odbor za umetno inteligenco (CAHAI). Njegov cilj je preučiti „izvedljivost in morebitne elemente pravnega okvira za razvoj, zasnovo in uporabo umetne inteligence, ki bi temeljili na standardih Sveta Evrope za človekove pravice, demokracijo in pravno državo"[42]. Aprila 2020 je Odbor ministrov Sveta Evrope sprejel priporočila o vplivu algoritemskih sistemov na človekove pravice[43].

Poleg tega je Organizacija za gospodarsko sodelovanje in razvoj (OECD) sprejela načela za umetno inteligenco in ustanovila opazovalno skupino za politiko umetne inteligence[44]. Na svetovni ravni UNESCO začenja z razvojem svetovnega instrumenta za določanje standardov na področju umetne inteligence[45]. To so le izbrani primeri širokega spektra pravnih in političnih pobud, katerih cilj je prispevati k določanju standardov na področju umetne inteligence. Vključeni so tudi dejanska zakonodaja (oz. osnutki zakonodaje), mehko pravo, smernice in priporočila za uporabo umetne inteligence ali poročila s priporočili za zakonodajo in politiko.

FRA je pripravila (neizčrpen) seznam pobud, povezanih z oblikovanjem politik za področje umetne inteligence[46]. Čeprav te vključujejo tudi zakonodajne pobude v državah članicah EU, so številne organizacije in podjetja sprožili pobude za odpravo etičnih pomislekov v zvezi z umetno inteligenco. Čeprav so etični pristopi koristni za reševanje morebitnih težav na področju umetne inteligence, se pogosto opirajo le na prostovoljne ukrepe. Na ta način ni v zadostni meri zagotovljeno spoštovanje temeljnih pravic.

Kot je FRA poudarila v svojem Poročilu o temeljnih pravicah za leto 2019: „Visoko raven zaščite pred morebitno zlorabo novih tehnologij in nepravilnostmi pri njihovi uporabi zagotavlja le pristop, ki temelji na pravicah.“[47] Pobuda Evropske komisije za regulacijo umetne inteligence pomaga preprečiti nepovezane odzive na umetno inteligenco v posameznih državah članicah, ki bi lahko bili škodljivi za podjetja v EU in subjekte zunaj EU.

Končne opombe

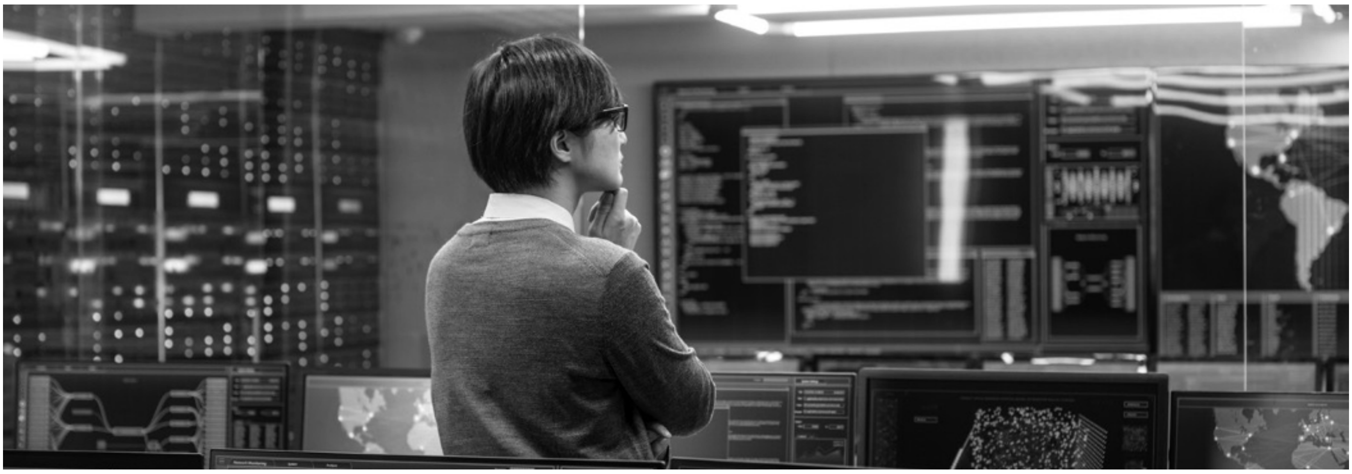
1 Agencija Reuters (2018), ***Amazon scraps secret AI recruiting tool that showed bias against women*** (Amazon umika skrivno orodje umetne inteligence za zaposlovanje, ki je pokazalo pristranskost do žensk), 10. oktober 2018.
2 Klepetalni boti sodijo med običajne funkcije umetne inteligence, vgrajene v aplikacije za pošiljanje sporočil, za simuliranje človeškega pogovora prek glasu ali besedila.
3 Independent (2017), ***AI robots learning racism, sexism and other prejudices from humans, study finds*** (Študija ugotavlja, da se roboti na podlagi umetne inteligence od ljudi učijo rasizma, seksizma in drugih predsodkov), 17. april 2017.
4 Prates, M., Avelar, P. in Lamb, L. (2019) ***Assessing Gender Bias in Machine Translation – A Case Study with Google Translate*** (Ocenjevanje razlikovanja na podlagi spola pri strojnem prevajanju – študija primera z uporabo prevajalnika Google), 11. marec 2019.
5 **Projekt Gender Shades, ki ocenjuje točnost produktov, ki uporabljajo umetno inteligenco za razvrščanje po spolu**.
6 Glej na primer: Der Standard (2020), ***Datenschutzbehörde kippt umstrittenen AMS-Algorithmus***, ali AlgorithmWatch (2019), ***Poland: Government to scrap controversial unemployment scoring system*** (Poljska: vlada umika sporni sistem za ocenjevanje brezposelnosti).

7 Privacy First (2020), *Dutch risk profiling system SyRI banned following court decision* (Nizozemski sistem oblikovanja profilov tveganja SyRI prepovedan s sodbo sodišča).
8 Evropska komisija (2020), *European enterprise survey on the use of technologies based on artificial intelligence* (Raziskava evropskih podjetij o uporabi tehnologij, ki temeljijo na umetni inteligenci), Luxembourg, julij 2020.
9 Glej npr. spletno mesto ***AI myths*** (Miti o umetni inteligenci).
10 Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F. in Delipetrev, B. (2020), ***AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*** (Opazovalnik umetne inteligence: opredelitev umetne inteligence: na poti k operativni opredelitvi in taksonomiji umetne inteligence), Luxembourg.
11 Schuett, J. (2019), *A legal definition of AI* (Pravna opredelitev umetne inteligence), **arXiv: 1909.01095**.
12 Hastie, T., Tibshirani, R. in Friedman, J. (2009), ***The Elements of Statistical Learning: Data Mining, Inference, and Prediction*** (Elementi statističnega učenja: podatkovno rudarjenje, poseganje in napovedovanje), Springer.
13 Glej na primer: Pasquale, F. (2015), *The Black Box Society. The Secret Algorithms That Control Money and Information* (Družba črne skrinjice: skrivni algoritmi, ki obvladujejo denar in informacije), Hardvard University Press, Cambridge in London; in Rai, A. (2020), „**Explainable AI: from black box to glass box**“ (Razložljiva umetna inteligenca: od črne do steklene skrinjice), *Journal of the Academy of Marketing Science*, zvezek 48, str. 137–141.
14 Temeljna razprava, ki opisuje to razliko, je: Breiman, L. (2001), „Statistical Modeling: The Two Cultures“ (Statistično modeliranje: dve kulturi), *Statistical Science*, 2001, zvezek 16, št. 3, str. 199–231.
15 **Resolucija Evropskega parlamenta z dne 14. marca 2017 o posledicah velepodatkov za temeljne pravice: zasebnost, varstvo podatkov, nediskriminacija, varnost in kazenski pregon** (2016/2225(INI)), odst. 1.
16 Prav tam, odst. 20.
17 Evropski svet (2017), **Zasedanje Evropskega sveta (19. oktober 2017) – sklepi**, EUCO 14/17, Bruselj, 19. oktober 2017, str. 8.
18 Evropska komisija (2018), **sporočilo Komisije Evropskemu parlamentu, Evropskemu svetu, Svetu, Evropskemu ekonomsko-socialnemu odboru in Odboru regij: Umetna inteligenca za Evropo**, COM(2018) 237 final, 25. april 2018.
19 Več informacij je na voljo na **spletni strani strokovne skupine na visoki ravni**.
20 Strokovna skupina na visoki ravni za umetno inteligenco (2019), Etične smernice za zaupanja vredno umetno inteligenco; **Priporočila glede politike in naložb za zaupanja vredno umetno inteligenco**.
21 Strokovna skupina na visoki ravni za umetno inteligenco (2020), **Ocenjevalni seznam za zaupanja vredno umetno inteligenco (ALTAI) za samoocenjevanje**.
22 Evropski svet, **Nova strateška agenda 2019–2024**, str. 4.
23 von der Leyen, U., **Bolj ambiciozna Unija. Moj načrt za Evropo**, str. 13.
24 Evropska komisija, *Bela knjiga o umetni inteligenci – evropski pristop k odličnosti in zaupanju*, COM(2020) 65 final, Bruselj, 19. februar 2020, str. 2.
25 Prav tam, str. 12.
26 Evropska komisija (2020), **Bela knjiga o umetni inteligenci: javno posvetovanje o evropskem pristopu k odličnosti in zaupanju**, 17. julij 2020.
27 Evropska komisija (2020), Prilagojeni delovni progrm Komisije 2020, Priloga I: Nove pobude, 27. maj 2020.
28 Evropski parlament, zakonodajni observatorij, **Okvir za etične vidike umetne inteligence, robotike in sorodne tehnologije**, 2020/2012(INL).
29 **Resolucija Evropskega parlamenta z dne 20. oktobra 2020 s priporočili Komisiji o ureditvi civilne odgovornosti za področje umetne inteligence**, 2020/2014(INL).
30 **Resolucija Evropskega parlamenta z dne 20. oktobra 2020 o pravicah intelektualne lastnine pri razvoju tehnologije umetne inteligence**, 2020/2015(INI).
31 Evropski parlament, **Umetna inteligenca v kazenskem pravu in njena uporaba v policiji in pravosodnih organih na področju kazenskih zadev**, 2020/2016(INI).
32 Evropski parlament, zakonodajni observatorij, **Umetna inteligenca v izobraževanju, kulturi in avdiovizualnem sektorju**, 2020/2017(INI).
33 **Sklep Evropskega parlamenta z dne 18. junija 2020 o ustanovitvi posebnega odbora za umetno inteligenco v digitalni dobi in opredelitvi njegovih pristojnostih, številčne sestave in mandata**, 2020/2684(RSO).
34 Evropski svet (2020), **Izredno zasedanje Evropskega sveta (1. in 2. oktober 2020) – sklepi**, EUCO 13/20, 2. oktober 2020.
35 Svet Evropske unije (2020), **Oblikovanje digitalne prihodnosti Evrope – sklepi Sveta**, 9. junij 2020.
36 Svet Evropske unije, **Sklepi Sveta „Dostop do sodnega varstva – izkoriščanje priložnosti, ki jih prinaša digitalizacija“**, 13. oktober 2020.
37 Svet Evropske unije, **Sklepi predsedstva – Listina o temeljnih pravicah v kontekstu umetne inteligence in digitalne preobrazbe**, 21. oktober 2020.
38 Glej npr. Pagallo, U., Casanovas, P. in Madelin, R. (2019), „**The middle-out approach: assessing models of legal governance in data protection, artificial intelligence, and the Web of Data**“ (Srednja pot pri pristopu k oceni modelov zakonitega upravljanja na področju varstva podatkov, umetne inteligence in spletnih podatkov), *The Theory and Practice of Legislation* 7 (1), str. 1–25.
39 Glej FRA (2019), ***Poročilo o temeljnih pravicah 2019,*** Luxembourg, Urad za publikacije, poglavje 7.
40 Sporočilo Komisije Evropskemu parlamentu, Svetu, Evropskemu ekonomsko-socialnemu odboru in Odboru regij, **Evropska strategija za podatke**, COM(2020) 66 final.
41 Evropski svet, **Sklepi izrednega zasedanja Evropskega sveta (17., 18., 19., 20. in 21. julij 2020)**, EUCO 10/20, 21. julij 2020.
42 Svet Evrope, *ad hoc* odbor za umetno inteligenco (CAHAI), **Informativno gradivo: Upravljanje digitalne preobrazbe**.
43 Svet Evrope, ***Priporočilo CM/Rec(2020)1 Odbora ministrov državam članicam o posledicah algoritemskih sistemov za človekove pravice*** (ki ga je Odbor ministrov sprejel 8. aprila 2020 na 1373. seji namestnikov ministrov).
44 Glej namensko **spletno mesto OECD**.
45 Glej namensko **spletno mesto Unesca**.
46 Glej pregled, ki ga je pripravila FRA, ***AI Policy Initiatives*** (Pobude politike za umetno inteligenco), ali pregled na **spletnem mestu Sveta Evrope**.
47 FRA (2019), ***Poročilo o temeljnih pravicah***, Luxembourg, Urad za publikacije, str. 166.

2.

UMESTITEV TEMELJNIH PRAVIC V USTREZEN KONTEKST – IZBRANI PRIMERI UPORABE UMETNE INTELIGENCE V EU

Uporaba tehnologij, povezanih z umetno inteligenco, je v EU razmeroma razširjena. Nedavna raziskava je pokazala, da 42 % podjetij že uporablja tehnologije, povezane z umetno inteligenco, 18 % podjetij pa njihovo uporabo načrtuje.



Pojasnilo o anketirancih

Primeri uporabe, predstavljeni v tem poglavju, temeljijo na informacijah, pridobljenih v razgovorih s predstavniki javnega in zasebnega sektorja.

Anketiranci iz javne uprave delajo na področju zdravstvenih storitev, infrastrukture in energetike, pravosodja, kazenskega pregona, migracij in upravljanja meja, socialnih prejemkov, davkov ter prometa in nadzora prometa.

Anketiranci iz zasebnih podjetij večinoma delajo v maloprodaji ali so aktivni na področju postavljanja cen ali trženja, zdravstva, finančnih storitev, energetike, zavarovanja, zaposlovanja, prometa ali na medsektorskih področjih s poudarkom na razvoju umetne inteligence za različne sektorje.

V tem poglavju so predstavljeni izbrani primeri, ko se uporablja umetna inteligenca – na področju umetne inteligence jih imenujemo primeri uporabe. FRA je informacije o takih primerih zbirala od petih držav članic EU: Estonije, Finske, Francije, Nizozemske in Španije. Vključena so različna področja uporabe v javni upravi in zasebnih podjetjih. Poseben poudarek je na uporabi umetne inteligence na področjih socialnih prejemkov, napovednega policijskega dela, zdravstvenih storitev in ciljno usmerjenega oglaševanja.

Poglavje vsebuje informacije o trenutni uporabi umetne inteligence in osnovne informacije o pristojnostih EU na teh izbranih področjih. S pomočjo teh primerov uporabe si lahko ustvarimo okvirno sliko o tem, katere vrste umetne inteligence in sorodnih tehnologij se trenutno uporabljajo.

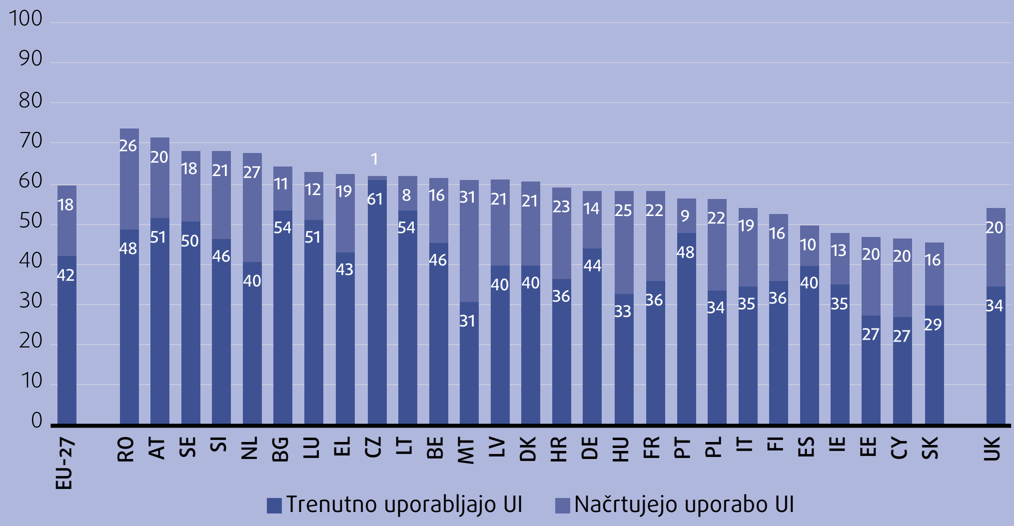
Primeri zagotavljajo tudi ustrezen kontekst za izvedbo analize z vidika temeljnih pravic. Preučitev širokega spektra primerov uporabe

Uporaba umetne inteligence s strani podjetij v EU v letu 2020

Po podatkih iz evropske raziskave podjetij je v začetku leta 2020 42 % podjetij v EU navedlo, da uporabljajo tehnologije, ki so odvisne od umetne inteligence. Ta odstotek sega od 27 % v Estoniji in na Cipru do 61 % na Češkem (glej sliko 1). Dodatnih 18 % podjetij pa uporabo umetne inteligence načrtuje v prihodnosti.

Raziskava kaže, da se umetna inteligenca uporablja predvsem v sektorju informacijske tehnologije (63 %). Uporabljene tehnologije vključujejo različne aplikacije informacijske tehnologije, namenjene optimizaciji procesov ali opreme, odkrivanju nepravilnosti, avtomatizaciji procesov in predvidevanju, optimizaciji cen in odločanju.

SLIKA 1: PODJETJA, KI SO V LETU 2020 UPORABLJALA UMETNO INTELIGENCO; PREGLED PO DRŽAVAH ČLANICAH (V %)



Opombe: V raziskavi smo spraševali glede uporabe ali načrtov za uporabo desetih različnih tehnologij, povezanih z umetno inteligenco, kot so prepoznavanje govora, vizualna diagnostika, odkrivanje goljufij, analiza čustev, predvidevanje na podlagi strojnega učenja in še več. Vključen je odstotek podjetij, ki so uporabljala vsaj eno od tehnologij umetne inteligence. N = 9 640.

Vir: FRA, 2020 [na podlagi podatkov, pridobljenih iz raziskave Evropske komisije, European enterprise survey on the use of technologies based on artificial intelligence *(Raziskava evropskih podjetij o uporabi tehnologij, ki temeljijo na umetni inteligenci), Luxembourg, julij 2020]*

zagotavlja pomemben vpogled v to, kako lahko dejanska uporaba umetne inteligence vpliva na temeljne pravice ljudi. **Poglavje 4** vsebuje razpravo o posledicah za temeljne pravice ob sklicevanju na primere, ki so v tem poglavju natančneje opisani.

Kot je bilo navedeno, se poročilo osredotoča na štiri splošne primere uporabe umetne inteligence:

- socialne prejemke,
- napovedno policijsko delo,
- zdravstvene storitve,
- ciljno usmerjeno oglaševanje.

Ta področja so z vidika temeljnih pravic še posebej občutljiva. Dve zajemata zlasti uporabo umetne inteligence v javni upravi (socialni prejemki in napovedno policijsko delo). Drugi dve se nanašata na zasebna podjetja (zdravstvene storitve in ciljno usmerjeno oglaševanje). Ti primeri uporabe predstavljajo podlago za analizo temeljnih pravic, ki je predstavljena v poročilu,

„Umetna inteligenca in strojno učenje sta različna pojma. Umetna inteligenca je krovni izraz.“
(zasebno podjetje, Estonija)

„Zdaj se dogaja, da vsakdo, ki uporablja strojno učenje, to označuje kot umetno inteligenco.“
(javna uprava, Nizozemska)

saj zagotavljajo potreben kontekst. Poročilo po potrebi izpostavlja tudi druge ugotovitve iz razgovorov, ki pokrivajo področja izven štirih omenjenih.

Na voljo so podrobne študije o taksonomiji umetne inteligence[1], ki zagotavljajo nadaljnjo kategorizacijo tehnologij. Kot je bilo ugotovljeno v uvodu, so imeli anketiranci različna stališča o tem, kaj umetna inteligenca je; nekateri so navedli, da jasne opredelitve umetne inteligence ni.

To poročilo obravnava posebne primere uporabe brez nadaljnjega razvrščanja uporabljene tehnologije. Vendar se je uporaba umetne inteligence v obravnavanih primerih razlikovala: uporaba tehnologije, ki so jo opisali anketiranci, je vključevala različne ravni kompleksnosti in različne ravni avtomatizacije.

Slika 2 vsebuje pregled različnih primerov uporabe, ki so jih anketiranci razumeli pod pojmom umetna inteligenca. Nekatere aplikacije so razmeroma preprosto razumljive. Pri sprejemanju odločitev, ki temeljijo na pravilih, so algoritmi opredeljeni na podlagi pravil „če-potem“ (na primer, če je dohodek določene osebe pod določenim pragom, potem bo ta oseba upravičena do določenih ugodnosti). Takšni algoritmi so bili uporabljeni na področju socialnih prejemkov na različnih ravneh avtomatizacije, pri čemer so bili vključeni primeri popolnega ali delnega človeškega nadzora in primeri odsotnosti takšnega pregleda.

Druge aplikacije so za generiranje odločitev uporabljale bolj tradicionalne statistične metode. Te so vključevale na primer *regresijsko analizo*. Gre za klasično statistično metodo, ki analizira korelacijo med več informacijami („spremenljivkami“) in izidom, ki je v tem primeru ocena kreditne sposobnosti. V drugih primerih so bile uporabljene bolj zapletene metodologije strojnega učenja, ki so vključene v pripravo napovedi in statistik za vladna poročila.

Obstajajo tudi algoritmi z veliko višjimi stopnjami kompleksnosti, kot je *globoko učenje* za diagnostično podporo na področju zdravja. Takšna orodja še vedno vključujejo visoko raven človeškega nadzora in posledično ne vključujejo visoke ravni avtomatizacije.

Nasprotno pa je ciljno usmerjeno oglaševanje primer potencialne uporabe zelo zapletenih algoritmov brez človeškega nadzora vseh rezultatov in odločitev, prav tako pa uporablja zelo zapletene algoritme, vključno z *globokim* in *spodbujevalnim učenjem* (za opise teh izrazov glej **poglavje 1**). Človeški nadzor na tem področju ne bi bil mogoč tudi zaradi obsega, v katerem taki algoritmi delujejo.

SLIKA 2: PRIMERI RAZLIČNIH RAVNI AVTOMATIZACIJE IN KOMPLEKSNOSTI V OBRAVNAVANIH PRIMERIH UPORABE

Avtomatizacija / Kompleksnost	nizka		visoka
visoka	avtomatizirano odločanje o socialnih prejemkih, ki temelji na pravilih	predvidevanje na podlagi regresijske analize za popolnoma avtomatizirano ocenjevanje kreditne sposobnosti	globoko učenje in spodbujevalno učenje za oglaševanje
	avtomatizirano odločanje v primeru pozitivnega izida postopka za dodelitev socialnih prejemkov, ki temelji na pravilih	predvidevanje na podlagi regresijske analize za ocenjevanje kreditne sposobnosti s človeškim nadzorom	tehnologija za prepoznavanje obraza za identifikacijo s človeškim nadzorom
nizka	človeško odločanje o socialnih prejemkih, ki temelji na pravilih	izdelava napovedi s podporo strojnega učenja	analiza medicinskih slik za diagnostične namene

Področja, kamor sodijo primeri:

- socialno varstvo
- trženje
- kazenski pregon
- zdravstvene storitve
- finančne storitve

Vir: FRA, 2020

Opombe: Primeri s področja finančnih storitev in uporaba tehnologije za prepoznavanje obraza niso zajeti v podrobnih opisih primerov uporabe, temveč so bili omenjeni v drugih razgovorih. Primeri ponazarjajo različne ravni kompleksnosti in avtomatizacije, kot se uporabljajo v praksi.

Sistemi umetne inteligence se razlikujejo tudi glede na morebitno škodo, ki bi lahko nastala zaradi napačne odločitve kot posledice uporabe umetne inteligence. Glede na konkretno področje uporabe imajo lahko napačne odločitve, ki so posledica napačnih rezultatov sistema, različne učinke. Pri uporabi umetne inteligence za odločanje so posledice različne, če je odločitev pritrdilna, vendar napačna (lažno pozitivna) ali negativna, vendar napačna (lažno negativna).

Ta vprašanja so še posebej pomembna pri uporabi strojnega učenja, ki temelji na statističnih izračunih, ti pa s seboj vedno prinašajo določeno stopnjo napake. Čeprav lahko algoritmi, ki temeljijo na pravilih, prav tako delajo napake (zlasti če postanejo kompleksnejši), so zaradi deterministične narave razvitih pravil tveganja v tem primeru manjša.

Na primer, če umetno inteligenco uporabljamo za sprejemanje odločitev o socialnih prejemkih, bo lažno pozitivna ocena pomenila, da oseba prejme dajatve, do katerih sicer ne bi bila upravičena. To ne pomeni nujno, da bo odločitev negativno vplivala na zadevno osebo (razen če se napaka ugotovi pozneje in je treba denar vračati). Vendar pa takšno stanje negativno vpliva na javno upravo, saj denar ni izplačan v skladu z dobro upravno prakso. Za razliko od tega pa bi lažno negativna odločitev negativno vplivala na posameznika, saj ne bi prejel ugodnosti, do katerih je upravičen. V Prilogi 2, ki je na voljo na **spletnem mestu FRA**, so navedeni hipotetični primeri učinkov napačnih odločitev, ki temeljijo na obravnavanih primerih uporabe.

Pomembno je, da bi se pri avtomatizaciji nalog učinki lahko povečali do mere, ki ima lahko negativne učinke na družbo kot celoto. Resnost in obseg morebitne škode sta samo dva od vidikov, ki jih je treba upoštevati pri analizi morebitnih posegov v temeljne pravice, ki jih s seboj prinaša uporaba umetne inteligence.

Majhna stopnja napak pri uporabi tehnologije za prepoznavanje obraza, ki jo uporabljajo organi kazenskega pregona, lahko denimo še vedno privede do označitve številnih nedolžnih ljudi, če se takšna tehnologija uporablja na krajih, kjer se analizira velika masa ljudi. To so lahko na primer letališča ali železniške postaje, kjer

Tehnologije, ki se uporabljajo v vseh primerih, opredeljenih v raziskavi

Anketiranci v večini primerov omenjajo **„strojno učenje“**, vključno z uporabo nevronskih mrež in njihovih razširitev (za opis strojnega učenja glej **poglavje 1**). Anketiranci so bodisi neposredno omenili dano področje ali so navedli podpodročja strojnega učenja, kot sta prepoznavanje slik ali tehnologija za prepoznavanje obraza (angl. *facial recognition technology*, FRT).

Najpogosteje so omenjali uporabo **„nadzorovanega strojnega učenja“**, ki se v glavnem uporablja za optimizacijo za posebej določen rezultat. Kljub temu se je v nekaterih primerih uporabljalo tudi **„nenadzorovano strojno učenje“** za kategorizacijo ali združevanje podatkov. Samo en primer se je nanašal na uporabo **„spodbujevalnega učenja“**, ne da bi se pri tem spuščali v podrobnosti.

Več anketirancev je omenilo uporabo **„obdelave naravnega jezika“** (angl. *natural language processing*, NLP). Gre za tehnologijo, namenjeno analizi besedila in govora, ki se včasih kombinira z algoritmi strojnega učenja.

Le redki omenjajo primere, ki vključujejo algoritme, ki temeljijo na pravilih, kar pomeni, da so pravila, ki jim algoritem sledi, neposredno kodirana (tj. temeljijo na pravilih „če-potem“).

V nekaterih primerih anketiranci niso razkrili ali niso mogli zagotoviti podrobnih informacij o uporabljeni tehnologiji.

bi lahko dnevno skenirali obraze več tisoč ljudi[2]. Morebitna pristranskost pri stopnji napak bi potem lahko povzročila nesorazmerno ciljno usmerjanje na nekatere skupine v družbi.

Praviloma so anketiranci omenjali več različnih primerov uporabe, vendar so bili pozvani, naj se med razgovorom osredotočijo na eno aplikacijo.

— Pomemben zaključek terenskega dela je tudi ugotovitev, da so podjetja in javne uprave pogosto še vedno čisto na začetku preučevanja uporabe umetne inteligence. Samo v dveh tretjinah primerov gre za dejansko uporabo v praksi. Številni primeri uporabe, ki so jih opisali anketiranci, so namreč v pilotni fazi, v razvoju ali še vedno v fazi raziskav.

SLIKA 3: BESEDE, KI SO JIH ANKETIRANCI NAJPOGOSTEJE UPORABILI PRI OPISOVANJU POSAMEZNIH PRIMEROV UPORABE

Opombe: Vizualizacija besed, ki so bile najpogosteje uporabljene pri opisih primerov uporabe, pripravljena s strani FRA. Večja kot je uporabljena pisava besede, pogosteje so anketiranci omenjali dani izraz.

Vir: FRA, 2020

- Dve aplikaciji, ki sta temeljili na umetni inteligenci, sta bili po opravljenem testiranju umaknjeni.

Slika 3 prikazuje najpogosteje uporabljene besede pri opisih primerov uporabe, ki so zajeti v tem poročilu. Poudarja pomen podatkov pri uporabi sistemov umetne inteligence in njihov pomen za podporo odločanju.

FRA je predhodno poudarila, da je natančen opis podatkov, ki jih uporabljajo aplikacije umetne inteligence, bistven za opredelitev in ublažitev morebitnih izzivov na področju temeljnih pravic[3]. Sistemi umetne inteligence, zajeti v tem poročilu, so uporabljali različne nabore podatkov. Vendar je bilo težko pridobiti podrobne informacije o uporabljenih podatkih, saj je večina anketirancev vire podatkov opredelila precej nejasno.

Na splošno je veliko anketirancev navedlo uporabo „odprtih podatkov“, „preteklih podatkov“ ali „metapodatkov“. Natančneje, anketiranci so omenili uporabo podatkov o strankah, npr. o nakupih ali vedenju pri spletnem brskanju, ali podatkov iz administrativnih evidenc, kot so podatki o socialnih prejemkih in davkih. Anketiranci so omenili tudi zdravstveno dokumentacijo, policijsko dokumentacijo, sodno dokumentacijo ter podatke iz družbenih medijev in podatke o prometu. Podatki so vključevali besedilne podatke (npr. e-pošto), avdio in video posnetke ter geolokacijske podatke. Podatki prihajajo iz notranjih zbirk podatkov podjetij in javne uprave, pa tudi iz zunanjih virov.

„Večinoma se uporablja, da prihranimo čas […], ko je treba pregledati velike količine gradiva.“
(javna uprava, Nizozemska)

Najpomembnejši razlog za uporabo umetne inteligence je povečana učinkovitost. Velika večina anketirancev iz javnega in zasebnega sektorja je omenila uporabo umetne inteligence zaradi hitrejšega postopka, manjšega števila napak in zmanjšanja stroškov, saj se potreba po človeških virih posledično zmanjša. Nekateri anketiranci iz vrst organov pregona so prav tako navedli, da uporabljajo umetno inteligenco za namene varnosti in zaščite ter za preprečevanje kaznivih dejanj.

„Najpomembnejša je učinkovitejša obravnava primerov. Gre za to, da čim učinkoviteje izkoristimo kapaciteto naših kadrov, ljudi, ki obravnavajo primere.“
(javna uprava, Nizozemska)

V preteklosti so naloge pri številnih primerih uporabe opravljali ljudje. Nekateri anketiranci so povedali, da uporabljajo umetno inteligenco, ker njena uporaba pomeni manj napak, kot če bi določene naloge opravljali ljudje. Nekateri anketiranci uporabljajo umetno inteligenco tudi za naloge, ki jih ljudje prej niso opravljali, saj takšnih količin informacij ljudje niso mogli obdelati – na primer na področju analize genoma ali prometnih napovedi.

Pomembno je, da je za približno polovico anketirancev uporaba umetne inteligence pomembna pri odločanju. Vendar se umetna inteligenca v glavnem uporablja zgolj za podporo odločanju, končne odločitve pa še vedno ostajajo večinoma v rokah ljudi.

Anketiranci so poudarili, da so tako javna uprava kot tudi zasebna podjetja pri uvajanju umetne inteligence kljub navdušenju še vedno previdni. Številni primeri uporabe so še vedno v fazi testiranja. Nekateri so bili v tej fazi tudi umaknjeni iz uporabe, kot je natančneje opisano v nadaljevanju. Kljub temu skoraj nihče od anketirancev ni vedel za kakršen koli načrt za zmanjšanje ravni uporabe tehnologije. V resnici je večina poudarila namere za vlaganje v inovacije ali nove načine uporabe trenutno razpoložljivih sistemov umetne inteligence.

2.1 PRIMERI UPORABE UMETNE INTELIGENCE V JAVNI UPRAVI

[Primer uporabe št. 1]

Avtomatizacija sistemov socialnega varstva – uporaba algoritmov na področju socialnih prejemkov

Ozadje in pravni okvir EU

Posebni poročevalec Združenih narodov za izredno revščino in človekove pravice Philip Alston je v svojem poročilu oktobra 2019 opozoril, da bi lahko uvedba „digitalne socialne države", ki bi vključevala uporabo umetne inteligence, privedla do „digitalne socialne distopije". Digitalizacijo sistemov socialnega varstva pogosto spremljajo znižanje skupnih proračunov za socialno varstvo, zmanjšanje števila upravičencev in drugi ukrepi, ki zmanjšujejo razpoložljivost sistemov socialnega varstva. Digitalizacija prav tako povečuje moč države, saj ponuja priložnosti za nadzor nad ljudmi. To je še posebej zaskrbljujoče v državah s precejšnjimi primanjkljaji na področju pravne države[4].

Uporaba algoritmov s strani javne uprave na področju socialne države vzbuja veliko zaskrbljenost zlasti zaradi njenega morebitnega negativnega vpliva na revščino in neenakost v primeru neustrezne uporabe na področju socialnih prejemkov[5]. To vključuje področja, kot so storitve socialnega varstva za otroke[6] in nadomestila za brezposelnost[7].

Kljub temu si javni organi prizadevajo za uporabo novih tehnologij, s katerimi bi odločanje o socialni varnosti in drugih prejemkih postalo učinkovitejše in potencialno pravičnejše. Na svetovni ravni se nove tehnologije uporabljajo za upravljanje sistemov socialne varnosti na več različnih načinov. Ti vključujejo preverjanje identitete, ocene upravičenosti, izračune višine prejemkov, preprečevanje in odkrivanje goljufij, točkovanje tveganja in klasifikacijo potreb ter komunikacijo med organi in upravičenci.

OECD socialne prejemke opredeljuje kot transferje gospodinjstvom, ki potrebujejo pomoč v zvezi z določenimi dogodki ali v posebnih okoliščinah, vključno z boleznijo, brezposelnostjo, upokojitvijo, nastanitvijo, izobraževanjem ali družinskimi razmerami[8]. Vendar splošno sprejete opredelitve socialnih prejemkov ni. Socialni prejemki, zlasti sistemi socialnega zavarovanja, se razlikujejo od sistemov zasebnega zavarovanja, saj vključujejo obvezne prispevke zaposlenih in delodajalcev, v nekaterih primerih v obliki obdavčitve[9].

Socialna politika, vključno s socialno varnostjo in socialno zaščito, sodi na področje deljene pristojnosti med EU in državami članicami (člen 4(2)(b) PDEU). V skladu s členom 151 PDEU si EU med drugim prizadeva za „izboljšanje življenjskih razmer in delovnih pogojev" ter „ustrezno socialno zaščito". V ta namen EU podpira in dopolnjuje dejavnosti držav članic na številnih področjih, vključno s socialno varnostjo in socialno zaščito delavcev ter bojem proti socialni izključenosti (člen 153(1) PDEU). EU lahko s svojimi ukrepi spodbudi sodelovanje med državami članicami in sprejme direktive z minimalnimi zahtevami. Poleg tega se odločitve na področju socialne varnosti in socialne zaščite lahko sprejemajo le po posebnem zakonodajnem postopku s soglasnim glasovanjem v Svetu[10].

Glede na povedano lahko države članice EU večino časa svobodno oblikujejo svoje politike socialne varnosti in socialne zaščite. Ker usklajevanja praktično ni, se sistemi socialne varnosti po vsej Uniji bistveno razlikujejo glede na konkretne dajatve, pogoje za upravičenost, način za izračun, plačevanje prispevkov in osebe, ki so jih dolžne plačevati, itd.

Javne uprave v državah članicah EU si prizadevajo za uvajanje umetne inteligence in sorodnih tehnologij na področju javne blaginje. Vendar pa so informacije o načinih uporabe teh tehnologij omejene. FRA je zbrala informacije o primerih uporabe, povezanih:

- z uporabo algoritmov pri podeljevanju nadomestil iskalcem zaposlitve,
- z obdelavo vlog za socialne prejemke in
- s strojnim učenjem podprto analizo podatkov o uporabi pokojnin.

Uporaba umetne inteligence v zasebnih zavarovalnicah

Več zasebnih zavarovalnic, s katerimi je bil opravljen razgovor v okviru te raziskave, uporablja umetno inteligenco in sorodne tehnologije. Vključena so področja obravnave zahtev strank za dopolnilno zdravstveno zavarovanje, podpore pri odločanju o nadomestilih iz naslova zavarovanja, ocenjevanja kreditnega tveganja posameznikov, oblikovanja cen zavarovanja, upravljanja zavarovalnih zahtevkov in podpore pri odločanju v zvezi s funkcijami upravljanja in kreditnih odločitev.

Zasebne zavarovalnice na splošno sprejemajo tehnologije, povezane z umetno inteligenco, saj te prispevajo k večji dobičkonosnosti njihovega poslovanja. Poročilo OECD poudarja pomen tehnologije v tem sektorju. Vendar pa izpostavlja, da bi klasifikacija tveganj lahko privedla do izključevanja tistih, ki sodijo v določeno ranljivo skupino, na način, ki je z družbenega ali političnega vidika nezaželen*.

* *OECD (2020),* The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector *(Vpliv masovnih podatkov in umetne inteligence (UI) v zavarovalniškem sektorju).*

Uporaba v praksi

Spodaj opisani primeri uporabe ponazarjajo nekatere izzive pri uporabi ali načrtovanju uporabe umetne inteligence na področju socialnih prejemkov v povezavi z algoritemskim odločanjem.

Eksperimentiranje z novimi tehnologijami pri podpori iskalcev zaposlitve

Javna organizacija je v okviru triletnega projekta eksperimentirala z več tehnologijami, povezanimi z umetno inteligenco, na področju svojega dela v povezavi z obdelavo nadomestil za iskalce zaposlitve in njihove podpore v smislu vrnitve na delo. Predstavnik, s katerim je bil opravljen razgovor, navaja, da lahko tehnologije, ki so predmet testiranja, izboljšajo in spodbudijo odnos z iskalci zaposlitve ter izboljšajo kakovost svetovanja tako v korist iskalcem zaposlitve kot tudi podjetjem. Po končanem testiranju se bo organizacija odločila, ali in kako bo te tehnologije uporabljala pri svojem vsakodnevnem delu.

Testiranje vključuje ugotavljanje privlačnosti ponudbe za zaposlitev, ki temelji na strojnem učenju, in sistem za ugotavljanje, ali iskalci zaposlitve še vedno aktivno iščejo zaposlitev.

Testiranje je vključevalo tudi profiliranje iskalcev zaposlitve z namenom ustreznega svetovanja. Vključen je bil izračun verjetnosti, da bo nekomu v določenem obdobju ponujeno prosto delovno mesto, in določitev parametrov, na podlagi katerih se izbere primerna ponudba zaposlitve. To se potem lahko odrazi v ustreznem svetovanju podjetjem o najboljših praksah za oblikovanje ponudb za zaposlitev. Profiliranje organizaciji omogoči določitev ustreznih storitev glede na profil in ozadje iskalca zaposlitve, kar nadomešča analize in sporočila, ki jih morajo pripravljati zaposleni. To bi v praksi potekalo tako, da bi od iskalcev zaposlitve zahtevali, da izpolnijo mesečni dnevnik o svojem iskanju zaposlitve. Vendar se še vedno preučuje, ali naj bo program omejen na zagotavljanje opisnih analiz ali naj gre še dlje in vključuje tudi priporočila. Organizacija glede tega vidika še vedno okleva.

Dodatno poteka testiranje sistema za obdelavo naravnega jezika za analizo vsebin e-poštnih sporočil iskalcev zaposlitve. Tukaj se elektronska pošta kategorizira, pridobijo se ustrezni podatki in ugotovita se nujnost in pomembnost e-poštnega sporočila. Proučuje se tudi možnost uporabe klepetalnega bota in samodejnih odgovorov na e-poštna sporočila.

Podatki, ki se uporabljajo za te sisteme, prihajajo iz več različnih virov znotraj organizacije. Podatki o iskalcih zaposlitve in njihovem ozadju, vključno z osebnimi davčnimi podatki, ter podatki o plačah in nadomestilih za socialno varnost se uporabljajo le pod zelo strogimi pogoji. Razlog leži v tem, da izvirajo iz virov podatkov, ki so podvrženi strogi regulaciji (npr. do plačilnih list sploh ni mogoče dostopati). Za pridobivanje znanja o trgu dela se uporabljajo tudi drugi podatki, kot so ponudbe podjetij za zaposlitev. Organizacija trenutno ne uporablja zunanjih podatkov, kot so podatki iz (profesionalnih) družbenih omrežij, ker uporaba takih podatkov ni urejena z nikakršnimi zakonskimi določbami.

Obdelava stanovanjskih ugodnosti – neuspeh in uspeh

Javni organ, pristojen za obdelavo socialnih prejemkov, je izvajal pilotni projekt, v okviru katerega je potekalo testiranje orodja umetne inteligence za obdelavo vlog in posledično podporo zaposlenim pri odločanju o stanovanjskih ugodnostih. Sistem je izbral primere novih vlog za ugodnosti, ki so bile sorazmerno enostavne za izračun. Te so vključevale nove vloge za stanovanjske ugodnosti, ki jih je vložil posameznik, ki živi sam ali z otroki, in posameznik, ki nima nobenih drugih prihodkov razen državnih ugodnosti. V splošnem so bili primeri ocenjeni kot preprosti, rezultat pa je bil, da je posameznik vedno prejel ugodnost.

Tehnološka rešitev je temeljila na modelu odločitvenega drevesa ob upoštevanju pravil za dodeljevanje stanovanjskih ugodnosti. Za izračun splošnih stanovanjskih ugodnosti je potrebna vnaprejšnja ocena dohodka. Podatki, uporabljeni med testiranjem, so izvirali iz interne baze podatkov, ki vsebuje podatke o postopkih vlaganja vlog za ugodnosti. Podatki so bili psevdonimizirani, ker ni bilo nobene potrebe po uporabi osebnih podatkov. Uporabljen je bil preprost statistični model (linearna regresija), pri čemer so bile vhodne spremenljivke dohodek in omejitve stroškov, izhodna spremenljivka pa znesek ugodnosti.

Vendar pa je bilo tudi v tako enostavnih primerih ugotovljeno, da je uporaba umetne inteligence v praksi prevelik zalogaj, razlog pa so bile pogoste spremembe zakonodaje. Posledično je bilo testiranje ustavljeno. Po mnenju anketiranca to, da ni pravne podlage za uporabo strojnega učenja, onemogoča njeno uporabo za sprejemanje upravnih odločb. Nadaljnjih načrtov za uporabo umetne inteligence za podporo odločanju o socialnih prejemkih ni.

Čeprav organizacija zaradi navedenih pravnih izzivov tega projekta ne izvaja več, je sogovornik prepoznal možnost nadaljnjih aplikacij in rešitev na tem področju v prihodnosti. Ugotovljeno je bilo, da so za organizacijo koristne zlasti umetna inteligenca in sorodne tehnologije, ki lahko podpirajo delovanje organizacije, ne da bi imele pravne učinke.

Hkrati organizacija uporablja obdelavo slik pri vlogah za socialne prejemke. Praviloma morajo prosilci za dajatve izpolniti več obrazcev in prilog, ki so pogosto predloženi v papirni obliki. Za učinkovitejše in časovno bolj ekonomično ravnanje z navedenimi dokumenti s strani zaposlenih v agenciji se prejete tiskane kopije skenirajo in nato razvrščajo ob uporabi avtomatiziranega sistema.

Prvi korak je, da se slike obrnejo v pravo smer. Algoritmi poravnajo dokumente, ki med skeniranjem niso bili pravilno poravnani, odstranijo madeže in očistijo ter uredijo barvo dokumenta, prepoznajo stolpce, odstavke, tabele in druge elemente kot ločene bloke, prepoznajo pisavo itd. Nato aplikacija preveri, ali sta prejeta vloga in priloga pravilno označeni (npr. če je dokument označen kot račun, sistem ugotovi, ali je oznaka pravilna).

Slike se obračajo in razvrščajo z uporabo tehnologije prepoznavanja slik in optičnega prepoznavanja znakov (angl. *optical character recognition*, OCR). Ta prepozna besedilo na slikah, vključno s fotografijami in skeni dokumentov ali ročno napisanih zapiskov. Tehnologija optičnega prepoznavanja znakov nato pretvori prepoznano besedilo v besedilne podatke, ki so strojno berljivi. Tu se v postopku prepoznavanja vzorcev najprej izolira vnos iz skeniranih slik, nato pa se primerja s t. i. glifi (tj. različicami črk), ki so shranjeni v sistemu na osnovi slikovnih pik.

Agencija bo nadaljevala z obdelavo slik in jo še naprej razvijala, na primer tako, da bo potencialno omogočila skeniranje črtnih kod iz prilog. To bi pospešilo potrjevanje pravilnosti dokumentov in prilog. Poleg tega bo na voljo tudi več rešitev, povezanih z obdelavo naravnega jezika.

Primer SyRI

Na Nizozemskem so razvili tako imenovani „sistemski kazalnik tveganja“ (angl. *system risk indication*, SyRI)* kot vladno orodje za obveščanje nizozemske javne uprave o tveganju goljufij državljanov prek obdelave in povezovanja velikih količin njihovih osebnih podatkov, ki so na voljo javnim organom.

Obsežna koalicija organizacij civilne družbe, ki se ukvarjajo z vprašanji zasebnosti, je sprožila tožbo, s katero je Okrožno sodišče v Haagu pozvala k preučitvi orodja SyRI, ki temelji na algoritmih**.

Sodišče je razsodilo, da SyRI nesorazmerno posega v zasebno življenje državljanov. Sodišče je ugotovilo, da so bili temu tveganju izpostavljeni vsi, katerih podatki so bili predmet analize v okviru tega orodja. Poleg tega zaradi nejasnosti uporabljenega algoritma državljani niso mogli „niti predvideti vdora v svojo zasebno življenje, niti se pred njim kakor koli zavarovati“***.

** Natančen opis orodja SyRI je na voljo v članku avtorja Ilje Brauna (2018),* **High risk citizens** *(Visoko tvegani državljani), v: Algorithm Watch.*

*** Sodba z dne 5. februarja 2020 (v nizozemščini) je na voljo na* ***spletu****.*

**** Privacy First (2020),* Dutch risk profiling system SyRI banned following court decision *(Nizozemski sistem oblikovanja profilov tveganja SyRI prepovedan s sodbo sodišča).*

Avtomatizacija nadomestil za brezposelnost

V eni od izbranih držav je večina odločitev o nadomestilih za brezposelnost popolnoma avtomatizirana. Nacionalna institucija, pristojna za nadomestila za primer brezposelnosti, je v letu 2019 posodobila svoj sistem v smislu popolne avtomatizacije večine obdelanih vlog za dajatve in odločbe. To je bilo storjeno po prilagoditvi ustrezne zakonodaje, ki je omogočila avtomatizirano sprejemanje odločitev.

Če se oseba prijavi kot brezposelna in vloži vlogo za nadomestilo, sistem pridobi informacije o prosilcu iz različnih drugih zbirk podatkov. To vključuje na primer register prebivalstva in zbirke podatkov davčnih organov, ki vsebujejo informacije o plačah in delovnih izkušnjah itd. Če so izpolnjeni vsi pogoji za prejemanje nadomestila za primer brezposelnosti, sistem izračuna obdobje izplačevanja na podlagi obdobja, v katerem je oseba vplačevala v sistem zavarovanja, in določi višino nadomestila na podlagi povprečne dnevne plače.

Postopek je popolnoma avtomatiziran. Vendar pa mora uslužbenec institucije posredovati, če iz zbirk podatkov ni mogoče pridobiti potrebnih informacij, če so v zbirkah podatkov navedene nasprotujoče si informacije ali če odločitev v konkretnem primeru vključuje določeno stopnjo diskrecije (tj. odločitve ni mogoče dokončno določiti na podlagi razpoložljivih podatkov in ima človek nekaj manevrskega prostora pri odločanju o primeru).

Glavni razlog za uporabo takšnega sistema je izboljšana učinkovitost. Poleg tega obstaja prepričanje, da sistem zagotavlja doslednost postopkov. To je zato, ker se vsaka vloga, ki ni predmet diskrecije, obravnava na enak način.

[Primer uporabe št. 2]

Napovedno policijsko delo – vnaprejšnje napovedovanje kaznivih dejanj

Ozadje in pravni okvir EU

Tehnologije umetne inteligence se uporabljajo na področju **kazenskega pregona**, zlasti pri napovednem policijskem delu. Obstoječe raziskave o tem, kako lahko takšna orodja vplivajo na temeljne pravice, so med drugim izpostavile posebna vprašanja v zvezi z diskriminacijo. Eden od ponavljajočih se pomislekov je možnost, da bi napovedno policijsko delo pomnožilo in utrdilo diskriminatorne prakse, zlasti z zanašanjem na pretekle podatke o kaznivih dejanjih, ki so lahko pristranski ali nepopolni. To je zato, ker številna kazniva dejanja, kot npr. nasilje v družini ali kazniva dejanja iz sovraštva, ostajajo v veliki meri neprijavljena in so zato v policijskih statistikah premalo upoštevana[11].

Poudarek na nekaterih kaznivih dejanjih, ki se zgodijo na javnih mestih, kot so na primer nasilje ali kazniva dejanja, povezana z drogami, za razliko od denimo poslovnih goljufij in neplačevanja davkov, lahko tudi zmanjša pravičnost odziva organov kazenskega pregona[12]. To je zato, ker so prvi pogosto povezani z določenimi demografskimi značilnostmi ali soseskami. V končni fazi bi bili na tak način ogroženi odnosi policije z nekaterimi skupnostmi.

Kriminološke raziskave o „žariščih kriminala“ potekajo že vrsto desetletij, zlasti v Združenem kraljestvu in ZDA[13]. Uporabljajo policijske podatke za kartiranje določenih kaznivih dejanj in izvajajo statistične teste za raziskovanje verjetnosti kaznivih dejanj. Različni policijski organi so jih uporabili in razvili v smislu obravnave različnih oblik koncentracij kaznivih dejanj („žarišč“).

V zadnjem času je bilo to področje uporabnih raziskav prilagojeno z uporabo umetne inteligence, ki služi kot orodje za izboljšanje učinkovitosti, nekateri pa nakazujejo, da bi uporaba algoritemskih orodij lahko zmanjšala zanašanje policije na subjektivno človeško presojo, ki lahko odraža pristranskost ali stereotipe[14]. Nekatere študije so prav tako pokazale, da bi lahko napovedno policijsko delo zmanjšalo nepotreben nadzor, zaslišovanje ter fizične preglede in preiskave[15], kar bi zmanjšalo število primerov poniževanja in nadlegovanja posameznikov, do katerih lahko pride med takšnimi aktivnostmi.

Napovedno policijsko delo je namenjeno napovedovanju verjetnosti kaznivih dejanj in predvidevanju nastajajočih trendov in vzorcev z namenom ustreznega prilagajanja strategij za preprečevanje kaznivih dejanj ali policijsko posredovanje[16]. Lahko je tudi del preiskave kaznivega dejanja, ki se je že zgodilo. Čeprav ni nobene splošno priznane opredelitve napovednega policijskega dela[17], je zanj značilna analiza podatkov za ugotavljanje skupnih vzorcev in trendov na področju kaznivih dejanj z uporabo algoritmov za oblikovanje modelov, ki temeljijo na tej analizi. To lahko služi za napovedovanje kriminalnih dejavnosti, ki bi se lahko pojavile v prihodnosti.

DEJAVNOST FRA

Preprečevanje nezakonitega profiliranja danes in v prihodnosti: priročnik

Pri razvoju in uporabi algoritemskega profiliranja se lahko v vsaki fazi postopka uvede pristranskost. Da bi preprečili to in morebitne poznejše kršitve temeljnih pravic, bi morali strokovnjaki za informacijsko tehnologijo in uradniki, ki razlagajo podatke, dobro poznati temeljne pravice.

Ta priročnik FRA ponazarja, kaj je profiliranje, predstavi pravne okvire, ki ga urejajo, in razlaga, zakaj je izvajanje profiliranja na zakonit način nujno za spoštovanje temeljnih pravic in ključno za učinkovito policijsko delo in upravljanje meja.

Za več informacij glej FRA (2018), **Preprečevanje nezakonitega profiliranja danes in v prihodnosti: priročnik**.

Namen tehnologij umetne inteligence na tem področju je na splošno bodisi „napovedati“ kazniva dejanja ali „napovedati“, kateri posamezniki bodo storili kazniva dejanja ali bodo žrtve kaznivih dejanj. Orodja za napovedovanje kaznivih dejanj se na splošno polnijo s preteklimi podatki – večinoma iz uradnih virov – o času, kraju in vrsti storjenih kaznivih dejanj. Ti so lahko dopolnjeni s spremenljivkami okolja, kot so gostota prebivalstva, obstoj nekaterih javnih prostorov ali storitev ter pomembni dogodki ali prazniki. Pri uporabi teh orodij se praviloma ne uporabljajo osebni podatki[18].

Nasprotno pa sistemi umetne inteligence, usmerjeni v napovedovanje morebitnih storilcev ali žrtev kaznivih dejanj, uporabljajo tako pretekle osebne podatke kot tudi podatke v realnem času. Ti bi lahko vključevali podatke iz kazenskih evidenc, naslove, telefonske številke, podatke o lokaciji, podatke, pridobljene iz družbenih medijev, informacije o znanih sodelavcih ter zdravstvene podatke ali podatke o dohodkih. Ti se nato združijo z drugimi podatki o kaznivih dejanjih in okolici[19].

EU in njene države članice si delijo pristojnosti na področju svobode, varnosti in pravice (člen 4(2)(j) PDEU). To vključuje tudi pravosodno sodelovanje v kazenskih zadevah in policijsko sodelovanje (členi 82–89 PDEU). Že ob sprejetju Lizbonske pogodbe je bilo v priloženi izjavi o varstvu osebnih podatkov na področju pravosodnega sodelovanja v kazenskih zadevah in policijskega sodelovanja ugotovljeno, „da bi lahko bila zaradi posebne narave [...] policijskega sodelovanja na teh področjih potrebna posebna pravila o varstvu osebnih podatkov in o prostem pretoku takih podatkov na podlagi člena 16 [PDEU]“[20].

V kontekstu napovednega policijskega dela so še posebej pomembni zbiranje, shranjevanje, obdelava, analiza in izmenjava informacij. Obdelava osebnih podatkov v okviru postopkov kazenskega pregona je na ravni EU urejena z direktivo o kazenskem pregonu (Direktiva (EU) 2016/680)[21]. Ta določa celovite standarde in zaščitne ukrepe za takšno obdelavo, vključno z varovanjem pred grožnjami za javno varnost in njihovim preprečevanjem.

Uporaba v praksi

Primeri uporabe, ki jih je zbrala FRA, kažejo na raznolikost načinov, na katere organi kazenskega pregona že uporabljajo ali načrtujejo uporabo umetne inteligence in sorodnih tehnologij za podporo svojemu delu.

Primeri, ki so jih omenili anketiranci, segajo od sistemov za podatkovno rudarjenje, ki so zasnovani za kartiranje vzorcev kaznivih dejanj, odkrivanja spletnega sovražnega govora in ocenjevanja tveganja za nasilje na podlagi spola do avtomatizacije nekaterih nalog paznikov v zaporih. Drugi primeri uporabe vključujejo odkrivanje nedovoljenih predmetov s satelitskih slik in, bolj na splošno, prepoznavanje predmetov na slikah. Poleg tega je bilo v raziskavah, opravljenih v zasebnem sektorju, omenjeno orodje za preprečevanje goljufij in odkrivanje kaznivih dejanj pri denarnih nakazilih.

Anketiranci so poudarili, da se umetna inteligenca in sorodni tehnološki sistemi uporabljajo za avtomatizacijo in pospešitev nalog, ki so jih prej opravljali ljudje, s čimer dosežemo sprostitev in/ali učinkovitejšo porazdelitev sredstev.

Kartiranje kaznivih dejanj v podporo učinkovitejši razporeditvi preiskovalnih zmogljivosti

Nacionalna obveščevalna agencija in urad javnega tožilca uporabljata sistem, ki temelji na podatkih, da bi svojim zaposlenim pomagala pri odločanju o tem, kako, kje in kdaj uporabiti razpoložljive preiskovalne zmogljivosti. Cilj je izboljšati razporeditev človeških virov in zagotoviti, da bodo uradne osebe prisotne ob pravem času na pravem mestu.

DEJAVNOST FRA

Tehnologija za prepoznavanje obraza v porastu: temeljni vidiki pravic v okviru kazenskega pregona

Zakonodaja EU med „občutljive podatke“ uvršča tudi podobo obraza, pri kateri gre v primeru obdelave s programsko opremo za prepoznavanje obraza za eno od vrst biometričnih podatkov. Vendar je takšne slike mogoče brez težav posneti tudi na javnih mestih. Čeprav se točnost ujemanja izboljšuje, tveganje za napake ostaja, in to zlasti pri nekaterih manjšinskih skupinah. Ljudje, katerih podoba obraza je zajeta in obdelana, morda ne vedo, da se to dogaja, in zato ne morejo nasprotovati morebitnim zlorabam.

Dokument FRA opisuje in analizira te in druge izzive v zvezi s temeljnimi pravicami, ki se pojavijo, ko javni organi uporabijo tehnologijo za prepoznavanje obraza za namene kazenskega pregona. Na kratko so predstavljeni tudi ukrepi, ki jih je treba sprejeti, da bi se izognili kršitvam pravic.

Za več informacij glej FRA (2019), **Facial recognition technology: fundamental rights considerations in the context of law enforcement** *(Tehnologija za prepoznavanje obraza: temeljni vidiki pravic v okviru kazenskega pregona).*

Anketiranci trdijo, da sistem lahko zagotovi natančnejše ocene v primerjavi z ljudmi, ki se pri odločanju pogosto zanašajo na svoj občutek. Kljub temu se sistem za sprejemanje operativnih odločitev vedno uporablja v kombinaciji s človeško presojo in drugimi sistemi, ki ne temeljijo na umetni inteligenci.

Na podlagi rezultatov, ki jih ustvari sistem, analitiki oblikujejo t. i. toplotni zemljevid. Ta ponazarja razširjenost nekaterih kaznivih dejanj na nekaterih območjih. Sistem je podoben ročni različici sistema za predvidevanje kaznivih dejanj, pri kateri so policisti z bucikami na zemljevidu označevali posebna območja tveganja. Uporabniki verjamejo, da lahko uporaba umetne inteligence celoten proces pospeši in poveča njegovo zanesljivost, saj lahko analizira večje količine podatkov.

Sistem temelji na rudarjenju podatkov in procesih strojnega učenja. Zasnovan je predvsem na edinstvenih policijskih podatkih iz poročil o kaznivih dejanjih ter izjav prič in osumljencev. Vrzeli se v največji možni meri rešujejo z uporabo drugih virov podatkov, kot so kriminološke raziskave ter socialni in demografski podatki, ki jih zagotovi nacionalni statistični urad. Sistem prav tako uporablja podatke iz odprtih virov.

Posebni parametri za izračun so odvisni od vrste kaznivega dejanja, saj se primernost napovednih dejavnikov na različnih področjih kaznivih dejanj razlikuje. V primeru vlomov se na primer zbirajo podatki o vlomih in združujejo s podatki o kraju prebivališča znanih prestopnikov in oddaljenosti od hiš, v katere je bilo vlomljeno. Pri tem se vnaprej določijo ustrezna merila, ki sistemu omogočijo izdelavo toplotnega zemljevida.

Napovedi na podlagi lokacije se pripravijo za naslednjih šest mesecev in vsebujejo podatek o času in lokaciji, na kateri bi lahko prišlo do vloma. Rezultat je zemljevid z majhnimi kvadratki, v katerih je tveganje za pojav kaznivih dejanj označeno z različnimi barvnimi odtenki. Anketiranci so navedli, da takšna vizualizacija policistom pomaga pri analizi sosesk in opazovanju korelacij med različnimi lokacijami.

Ocena tveganja za nasilje na podlagi spola v družini

Nacionalna policija uporablja interni sistem za spremljanje primerov nasilja na podlagi spola v družini. Sistem pomaga policistom pri sprejemanju odločitev in razporejanju sredstev v primerih nasilja v družini. Sistem kategorizira primere na podlagi ocenjenega ponovitvenega tveganja z namenom osredotočanja na najbolj tvegane primere.

Strokovna ekipa bi analizo tveganja lahko opravila tudi brez uporabe umetne inteligence. Vendar pa je sistem sposoben obdelati velike količine podatkov v kratkem času in pomagati manj izkušenim in nespecializiranim policistom pri analizi tveganja.

Ko se pojavi poročilo o primeru domnevnega nasilja na podlagi spola v družini, policist uvede preiskavo. Ta vključuje zbiranje dokazov, zasliševanje prič in morebitno prijetje. Na podlagi informacij, zbranih v tem postopku, policist izpolni podrobna vprašalnika za oceno navedb in verjetnosti ponovne kršitve, proučitev razvoja primera in oceno vedenja storilca in žrtve. Policist prav tako navede tudi stopnjo resnosti, naravo groženj in odnos do žrtve.

Sistem nato določi oceno tveganja ob upoštevanju tristopenjske lestvice. Policist lahko ročno poveča stopnjo tveganja, a ne more znižati ravni tveganja pod stopnjo, ki jo generira sistem. Ko je raven potrjena, se uporabijo posebni ukrepi v skladu z uveljavljenimi policijskimi protokoli. Sistem prek avtomatiziranega sistema tudi obvešča sodnika o potencialno hujših primerih.

DEJAVNOST FRA

Zaznavanje sovražnega govora na spletu

Javna agencija za boj proti kaznivim dejanjem iz sovraštva uporablja orodje, ki temelji na umetni inteligenci, za odkrivanje sovražnega govora na spletu prek analize vzorcev govora na spletu. Na podlagi obdelanih podatkov sistem določi, katere družbene skupine so potencialne žrtve. Na podlagi tega lahko organi kazenskega pregona sprejmejo ukrepe za njihovo zaščito, preden se grožnje uresničijo.

Čeprav je namen orodja identifikacija morebitnih žrtev in ne storilcev, lahko organi pregona uporabijo informacije, ki jih generira sistem, da od ponudnikov družbenih medijev zahtevajo informacije o posameznih uporabnikih za namene kazenske preiskave.

Poseben izziv je razumevanje konteksta, v katerem so bile izjave podane. Novinarji ali akademiki lahko na primer uporabljajo besede, povezane s sovražnim govorom, za poročanje ali analizo danega pojava.

V letu 2021 je nameravala FRA začeti z raziskavami o spletni prisotnosti sovraštva v družbenih medijih. Na ta način bo FRA lahko prispevala k razvoju politike na področju moderiranja spletnih vsebin z uporabo umetne inteligence.

2.2 PRIMERI UPORABE UMETNE INTELIGENCE V ZASEBNEM SEKTORJU

[Primer uporabe št. 3]

Umetna inteligenca in zdravje – analiza zdravstvene dokumentacije za reševanje življenj

Ozadje in pravni okvir EU

Področje zdravstvenega varstva ima v razpravah o uporabi umetne inteligence še posebej pomembno vlogo. Zdravstveni podatki in spletne aplikacije lahko izboljšajo rezultate zdravljenja in posledično prinašajo širše družbeno-gospodarske koristi. Pandemija covida-19 je še dodatno okrepila proaktivnost in zanimanje na tem področju, zlasti v smislu možnosti za (spletne) podatke in aplikacije za povečanje sposobnosti vlad in zdravstvenih služb za spremljanje širjenja bolezni.

Zdravstvo ima pomembno vlogo tudi v pogledih splošnega prebivalstva na uporabo umetne inteligence. Raziskava Eurobarometra iz leta 2019 je pokazala, da vsak drugi Evropejec meni, da se lahko umetna inteligenca najbolje uporabi za izboljšanje medicinske diagnostike, razvoj prilagojene medicine ali izboljšanje kirurških posegov[22].

Ta primer uporabe zajema uporabo umetne inteligence ali sorodnih tehnologij s strani deležnikov javnega in zasebnega sektorja na področju zdravstvenih evidenc in napovedovanja bolezni. Podatke iz elektronskih zdravstvenih kartonov (angl. *electronic medical records*, EMR) in elektronskih zdravstvenih zapisov (angl. *electronic health records*, EHR) je mogoče vnašati v sisteme umetne inteligence in sorodnih tehnologij, ki so namenjeni podpiranju razvoja preventivne medicine, ki omogoča prepoznavanje zgodnjih tveganj bolezni in določanje ustreznih posegov. Raziskovalci lahko predvidijo klinične dogodke, kot so smrtnost, hospitalizacija, ponovni sprejem in dolžina bivanja v bolnišnici.

Poleg napovedovanja bolezni je mogoče analizirati podatke iz zdravstvene kartoteke, da se predvidi pacientovo upoštevanje navodil pri zdravljenju in njegovo spoštovanje urnika zdravstvenih pregledov. Te tehnologije lahko podprejo boljše rezultate zdravljenja ter izboljšajo učinkovitost zdravstvenega sistema.

V skladu s členom 6 PDEU ima Evropska unija podporno pristojnost za varovanje in izboljšanje človekovega zdravja. Države članice so še naprej v celoti odgovorne za določanje svojih politik na področju zdravstva, organizacijo in upravljanje svojih zdravstvenih sistemov ter zagotavljanje zdravstvenih storitev (člen 168(7) PDEU).

Znotraj njene pristojnosti je dejavnost Evropske unije, ki dopolnjuje nacionalne politike, usmerjena k izboljševanju javnega zdravja, preprečevanju telesnih in duševnih obolenj in bolezni ter odpravljanju vzrokov, ki ogrožajo zdravje ljudi. Takšni ukrepi lahko zajemajo informiranje in izobraževanje na področju zdravja ter spremljanje, zgodnje opozarjanje in boj proti resni čezmejni ogroženosti zdravja (člen 168(1) PDEU). Na teh področjih lahko EU sprejme spodbujevalne ukrepe, ki ne vključujejo nikakršnega usklajevanja zakonov ali drugih predpisov držav članic.

Cilj drugih pravil in politik, sprejetih na ravni EU, je zagotoviti prost pretok državljanov, njihovo enako obravnavanje in varstvo pred diskriminacijo v tujini ter razpoložljivost in varnost medicinskih izdelkov in storitev na enotnem trgu. Upoštevanje razvoja tehnologij in njihove uporabe v zdravstvu, izmenjava zdravstvene dokumentacije, pravice pacientov v čezmejnih situacijah in napovedovanje bolezni so z vidika javnega zdravja še posebej pomembni.

V skladu z GDPR so zdravstveni in genski podatki opredeljeni kot posebna vrsta podatkov (člen 9), ki jih imenujemo „občutljivi podatki“[23]. Zahtevajo posebno zaščito, saj bi njihova obdelava lahko privedla do resnih tveganj. Zdravstveni in genski podatki posameznikov, na katere se nanašajo osebni podatki, se lahko izmenjujejo le pod posebnimi pogoji, ki jih določa člen 9(2) GDPR. GDPR zagotavlja izjemo od načela omejitve namena, če se podatki uporabljajo za raziskovalne namene v skladu s členom 89(1). Raziskovalci morajo zagotoviti, da so pri uporabi podatkov o pacientih vzpostavljeni tehnični in organizacijski zaščitni ukrepi, kot sta psevdonimizacija in anonimnost.

EU je sprejela tudi ukrepe v zvezi z izmenjavo zdravstvene dokumentacije. Priporočilo Evropske komisije C(2019)800 o evropski obliki izmenjave elektronskih zdravstvenih zapisov[24] si „prizadeva olajšati čezmejno interoperabilnost elektronskih zdravstvenih zapisov v EU s podporo državam članicam pri njihovih prizadevanjih za zagotovitev, da lahko državljani varno dostopajo do svojih zdravstvenih podatkov in si jih izmenjujejo ne glede na to, kje v EU se nahajajo“[25]. Priporočilo določa tehnične specifikacije za izmenjavo teh podatkov med državami članicami EU.

Tudi evropska strategija za podatke (februar 2020) je izrazito osredotočena na zdravstvene podatke[26]. „Skupni evropski zdravstveni podatkovni prostor“ je eden od devetih skupnih evropskih podatkovnih prostorov, katerih vzpostavitev bo podprla Evropska komisija.

Sistem zgodnjega obveščanja in odzivanja (angl. *Early Warning and Response System*, EWRS) je v lasti Evropske komisije in ga upravlja Evropski center za preprečevanje in obvladovanje bolezni. Njegova cilja sta „obveščanje na ravni EU o resni čezmejni ogroženosti zdravja“[27] in omogočanje „stalnega obveščanja Evropske komisije in držav članic EU za namene opozarjanja, ocenjevanja tveganj za javno zdravje in določanja ukrepov, ki so morda potrebni za varovanje javnega zdravja“[28].

Elektronski zdravstveni karton (EMR), ki je računalniško podprt zdravstveni karton, ustvarjen za paciente zdravstvene organizacije[29], in elektronski zdravstveni zapis (EHR), ki vsebuje pacientovo zdravstveno anamnezo, ki presega eno organizacijo in vključuje izmenjavo podatkov znotraj zdravstvenega sistema, lahko vključujeta velike količine osebnih podatkov. Ti lahko med drugim zajemajo: ime in podatke za stik posameznika in njegovih najbližjih sorodnikov, demografske informacije, diagnoze in rezultate testov ter zdravila in zdravljenje[30]. Vključujejo lahko tudi podatke o pacientih, ki so pridobljeni iz nosljivih pripomočkov[31].

Države članice Evropske unije nimajo enotnega sistema EMR/EHR[32]. Nekatere države, kot je Nemčija, sploh nimajo nacionalnega sistema EMR/EHR. Druge, vključno z Belgijo in Dansko, imajo različne sisteme EMR/EHR na regionalni ravni. Sistemi se zelo razlikujejo glede na to, kateri podatki se beležijo in kdo ima dostop do katerih podatkov[33]. Evropska komisija in drugi deležniki so kot glavno oviro za vzpostavitev enotnega digitalnega zdravstvenega trga izpostavili raznolikost sistemov EMR/EHR na ravni držav in pomanjkanje interoperabilnosti[34].

Študije poudarjajo potencial umetne inteligence ali sorodnih tehnologij za zgodnejšo diagnozo, širitev možnosti za preprečevanje bolezni in izboljšanje varnosti pacientov[35], kar bi okrepilo pravico dostopa do preventivnega zdravstvenega varstva in povečalo koristi zdravljenja. EMR/EHR bi lahko pripomogla tudi k večji personalizaciji zdravstvenega varstva[36], medtem ko bi lahko možnost hitre izmenjave podatkov zagotovila bolj usklajeno in pravočasno zdravljenje.

Vendar uporaba EMR/EHR predstavlja veliko tveganje z vidika varstva podatkov. Zdravstveni sektor je vodilni v smislu kršitev varnosti osebnih podatkov[37]. Zaradi velike količine shranjenih osebnih podatkov, ki je največja med vsemi panogami, skupaj s široko mrežo za izmenjavo podatkov in ogromnim številom točk dostopa je zdravstveni sektor privlačen cilj za hekerje[38].

Nekaj pomislekov vzbuja tudi kakovost podatkov v EMR/EHR. Študije, v katerih so pacientom pokazali njihove zdravstvene kartoteke in jih povprašali o njihovi točnosti, so pokazale, da je bilo do 50 % informacij nepopolnih ali napačnih[39]. Veliko pomembnih podatkov v EMR/EHR je nestrukturiranih in predstavljenih v obliki prostega besedila, kar dodatno zmanjšuje kakovost podatkov[40]. Nizka natančnost, nepopolnost in pomanjkljiva splošna kakovost podatkov povečujejo tveganje za pojav medicinskih napak[41].

Uporaba v praksi

Aplikacije, opisane v razgovorih, vključujejo preproste in naprednejše modele, ki se uporabljajo v **javnem in zasebnem sektorju**. Največ primerov uporabe se nanaša na diagnostična orodja na podlagi slik. Vendar pa so anketiranci govorili tudi o orodjih za avtomatizacijo različnih delovnih postopkov, kot so npr. kartiranje besedilnih podatkov, evidentiranje zdravstvene dokumentacije ter analize in meritve telesnih tkiv in živčnih vlaken.

Manjše število primerov se je nanašalo na naprednejše projekte, kot so sistemi za daljinsko spremljanje nekaterih zdravstvenih kazalnikov, kot je srčni utrip. V vsakem primeru sistemi dopolnjujejo strokovno znanje zdravstvenih delavcev. V naslednjih poglavjih so predstavljeni primeri diagnostičnih orodij in orodij za spremljanje na daljavo.

Uporaba umetne inteligence za ciljno usmerjanje zdravstvenih inšpekcij

Javni organ, pristojen za pregledovanje standardov varnosti hrane v restavracijah, uporablja strojno učenje za obdelavo podatkov o ocenah strank z večjih spletnih platform. Ti pomagajo pri odločitvi, kje in kdaj izvajati inšpekcijski nadzor. Prej je ta postopek temeljil na prejetih pritožbah in prejšnjih poročilih. Od uvedbe orodja se je stopnja restavracij, pri katerih je bilo ugotovljeno neskladje, podvojila s približno 18 % na 36 %.

Prvi korak vključuje besedilno rudarjenje. Algoritem identificira ocene, ki vsebujejo ključne besede, ki lahko kažejo na zdravstvene težave in varnostna vprašanja, kot npr. „zbolel", „slabost" ali „glodavci". V drugem koraku organ primerja rezultate, ki izhajajo iz ocen strank, s prejšnjimi poročili o inšpekcijskih pregledih z namenom izboljšanja točnosti in zanesljivosti algoritma.

Orodja na podlagi slik za pomoč pri odkrivanju in diagnosticiranju bolezni

Vsa orodja, ki se uporabljajo za podporo odkrivanju in diagnosticiranju bolezni, ki jih opisujejo anketiranci, delujejo na podoben način. Bolnišnica v zasebni lasti na primer uporablja sistem umetne inteligence za interpretacijo CT-posnetkov pacientov z možgansko kapjo. Po možganski kapi se uporabi slikanje za odkrivanje, kje je prišlo do poškodbe možganov in kje lahko pride do prekinitve preskrbe možganov s krvjo. Omogoča tudi meritve, ki jih zdravnik lahko primerja z določenimi vrednostmi.

Anketiranec meni, da aplikacija pomaga hitreje razbrati takšne značilnosti na sliki, kar lahko izboljša kakovost diagnoze, kar pa je odvisno tudi od tega, kdo orodje uporablja. Obenem poudari, da zanašanje na aplikacijo umetne inteligence ni nujno učinkovitejše, saj mora biti zdravstveni delavec, ki bi lahko sliko preučil, v vsakem primeru prisoten. Orodje pa lahko ponudi dodatno podporo – na primer, če ima zdravnik težave pri interpretiranju določene slike oziroma iskanju nepravilnosti.

Sistem je bil zgrajen, usposobljen in validiran z uporabo podatkovnega niza, ki je delno temeljil na obsežni znanstveni študiji, h kateri je bolnišnica prispevala. To je bilo dopolnjeno z nakupom tujih podatkovnih nizov. Algoritem v prihodnosti ne bo dodatno usposobljen ali prilagojen na podlagi novih podatkov. Prav tako ne bodo izdane nove različice. Razvijalci menijo, da če sistemu dovolimo, da se še naprej uči, bomo imeli težave pri validaciji njegovega delovanja.

Zasebno podjetje je razvilo algoritem, ki omogoča odkrivanje raka dojke iz izvidov mamografije. Orodje omogoča tolikšno verjetnost in stopnjo gotovosti, da lahko radiologom pomaga, da pospešijo analizo rezultatov in se odločijo, ali so upravičeni dodatni pregledi. Algoritem zaznava in opredeljuje anomalije v izvidu mamografije kot rakave ali nerakave.

Medtem ko anketiranec navaja, da ima sistem trenutno zelo nizko stopnjo lažno negativnih ali lažno pozitivnih rezultatov, ugotavlja, da v mnogih primerih ne zagotovi jasnega izvida. Sistem je bil usposobljen na podlagi podatkov iz radiografskih in mamografskih pregledov na območju Evrope in EU, pri čemer so bila pisna poročila in pretekle biopsije uporabljeni kot kontrolni podatki.

Spremljanje statističnih podatkov o vitalnih znakih pacientov na daljavo

Bolnišnica vodi pilotni projekt, v okviru katerega preizkuša sistem, ki podpira zgodnje odkrivanje potencialnih bolezni. Spremljanje kazalnikov zdravja pacientov (npr. krvnega tlaka ali srčnega utripa) praviloma poteka ročno in zajema situacijo v določenem trenutku. Nenehno spremljanje takšnih kazalnikov ima potencial za prepoznavanje trendov, ki jih zdravniki sicer ne bi prepoznali, in zgodnje zaznavanje težav z namenom preprečevanja bolezni. Sistem uporablja biosenzor (kot nekakšen obliž), ki beleži hemodinamične podatke pacientov prek nenehnega spremljanja srčnega utripa in dihanja.

Podatki, ki jih sistem uporablja, prihajajo iz bolnišnice in od pacienta. Ti podatki so anonimizirani in šele nato se izmenjujejo s tretjim ponudnikom. Razen informacij, zbranih na podlagi spremljanja s pomočjo obliža, se za izgradnjo in usposabljanje sistema ne uporabljajo nobene druge informacije. Podatki o dejavnikih v okolici niso bili vključeni v pilotni projekt, ker bi po mnenju anketiranca lahko pomenili pristranskost.

V prihodnosti bo sistem združil informacije, ki jih bo zbral biosenzor, z ločenimi informacijami iz EMR pacientov z namenom oblikovanja sklepov na podlagi trendov, opaženih pri spremljanju.

[Primer uporabe št. 4]

Ciljno usmerjeno oglaševanje – profiliranje potrošnikov za povečanje dobička

Ozadje in pravni okvir EU

Internet je spremenil način življenja. Veliko ljudi dnevno uporablja internetne storitve, ki so pogosto na voljo brezplačno. Podjetja, ki ponujajo svoje storitve brezplačno, svoj prihodek ustvarijo v pretežni meri z oglaševanjem, oglasi pa so samodejno usmerjeni k posameznim potrošnikom na podlagi informacij o njih.

Razpoložljivost podatkov o spletnem obnašanju posameznikov v kombinaciji s tehnologijami strojnega učenja je znatno izboljšala sposobnost trgovskih podjetij, da se ciljno usmerijo na določene posameznike. To bi lahko segalo celo do manipulacije potrošnikov prek predvidevanja njihovega odziva, ki temelji na nerazumskih vidikih psihologije in ne na utemeljeni izbiri[42].

Škandal s Cambridge Analytico je izpostavil zlasti negativen vpliv takšnih načinov uporabe v politične namene. V tem primeru je podjetje nezakonito pridobilo osebne podatke milijonov uporabnikov družbenih medijev za namene ciljnega usmerjanja političnih oglasov k različnim družbenim skupinam na podlagi določenih psiholoških profilov[43].

Nedavna izjava Odbora ministrov Sveta Evrope poudarja pomanjkanje znanja o manipulativni moči algoritmov. „Učinki ciljno usmerjene uporabe nenehno naraščajočih količin zbirnih podatkov o uveljavljanju človekovih pravic v širšem smislu, ki znatno presegajo sedanje koncepte varstva osebnih podatkov in zasebnosti, še vedno ostajajo premalo raziskani in jih je treba natančno preučiti."[44] Prav tako so se pojavili pomisleki glede tega, kako lahko spletno oglaševanje, ki temelji na tehnologijah umetne inteligence, vpliva na varstvo podatkov in zasebnost[45], varstvo potrošnikov[46], pravico do nediskriminacije[47] in celo način delovanja demokracij[48].

Beseda „oglaševanje" se nanaša na sporočila, namenjena vplivanju na vedenje potrošnikov. Oglaševanje v eni ali drugi obliki je bilo vedno usmerjeno k določenim skupinam, ki so bile izbrane na podlagi določenih značilnosti ali vedenja[49].

Porast družbenih medijev pa je ciljno usmerjeno oglaševanje dvignil na višjo raven zaradi neposrednega dostopa do podatkov o potrošnikih. Mikrociljanje pomeni usmerjanje v zelo specifične skupine – in več kot je podatkov, zbranih prek spletnih aktivnosti, bolj ciljno usmerjene so lahko te aktivnosti. Ker ponudniki družbenih medijev in platforme, kot sta Google ali Amazon, zbirajo celovite uporabniške podatke s spremljanjem različnih aktivnosti svojih uporabnikov, imajo oglaševalci dostop do vse podrobnejših informacij[50].

Področje ciljno usmerjenega oglaševanja in sistemov za priporočanje vsebin (npr. novic ali filmov) je eden redkih primerov, ki vključuje uporabo tako imenovanega spodbujevalnega učenja v resničnem življenju. Gre za tehnologijo, ki temelji na optimizaciji določenega cilja z eksperimentiranjem in samodejnim posodabljanjem pravil za doseganje čim boljših možnih rezultatov. To pomeni, da sistem preizkuša nameščanje različnih oglasov po načelu poskusov in napak in na ta način išče najboljši način za optimizacijo prihodkov, pri čemer je vključen element samoučenja.

Medtem ko imajo evropske države o dejanski uporabi spodbujevalnega učenja zelo omejeno znanje, velika podjetja, ki so aktivna na tem področju, to vprašanje raziskujejo[51].

Vprašanja, povezana s ciljno usmerjenim oglaševanjem, sodijo na področje varstva potrošnikov. To pa v skladu s členom 4(2)(f) PDEU sodi na področje deljene pristojnosti EU in držav članic. Ukrepi EU za varstvo potrošnikov so namenjeni varovanju zdravja, varnosti in ekonomskih interesov potrošnikov, pa tudi spodbujanju njihove pravice do obveščenosti, izobraževanja in samoorganiziranja za zaščito njihovih interesov (člen 169(1) PDEU). EU lahko sprejme minimalne usklajevalne ukrepe za zagotavljanje visoke ravni varstva potrošnikov (člen 114(3) PDEU), vendar državam članicam EU omogoča, da na nacionalni ravni uvedejo še strožje ukrepe.

V sekundarni zakonodaji EU pravila oglaševanja ureja Direktiva 2006/114/ES o zavajajočem in primerjalnem oglaševanju[52]. Ta direktiva zagotavlja minimalno raven varstva pred zavajajočim oglaševanjem. Usklajuje tudi pravila za primerjalno oglaševanje po vsej Uniji. Določbe Direktive 2006/114/ES se uporabljajo za odnose med potrošniki in podjetji ter odnose med podjetji. Vendar se praktično uporabljajo samo za slednje[53], odkar je začela veljati Direktiva 2005/29/ES o nepoštenih poslovnih praksah podjetij v razmerju do potrošnikov na notranjem trgu[54].

Storitve, ki vključujejo oglaševanje, dodatno pokriva tudi Direktiva 2006/123/ES o storitvah na notranjem trgu[55]. Poleg tega se uporablja tudi Direktiva 2000/31/ES o nekaterih pravnih vidikih storitev informacijske družbe, zlasti elektronskega poslovanja na notranjem trgu (direktiva o elektronskem poslovanju). Ta direktiva je del pravnega okvira EU za digitalne storitve. Da bi lahko zadovoljili potrebe na področju novih spletnih storitev in praks, je direktiva o elektronskem poslovanju trenutno v postopku revidiranja v okviru svežnja o digitalnih storitvah. Cilj tega svežnja je „okrepiti enotni trg digitalnih storitev ter spodbujati inovacije in konkurenčnost evropskega spletnega okolja“[56].

FRA je zbrala informacije o dejanskih primerih uporabe od šestih evropskih podjetij, ki se ukvarjajo z nameščanjem spletnih oglasov, priporočili glede vsebine in personaliziranim trženjem.

Uporaba v praksi

Zajeti primeri vključujejo:

- nameščanje spletnih oglasov na podlagi napovedovanja klikov (tj. raziskovanje verjetnosti, da spletni uporabniki kliknejo na določene povezave ali oglase) in avtomatizirano oddajanje ponudb na dražbah za spletni oglasni prostor,
- personalizirano in ciljno usmerjeno trženje in komuniciranje po e-pošti.

Večina nalog je bila popolnoma avtomatiziranih. Primeri se nanašajo na analize uporabniških preferenc in aktivnosti uporabnikov ter izračune verjetnosti klikov in nakupov, vključno z merjenjem učinkovitosti predhodno podanih priporočil. Vključene so tudi metode ciljno usmerjene komunikacije na podlagi opredeljenih ciljnih skupin za vzpostavitev (dolgoročnega) zaupanja med strankami in ponudniki storitev.

Ciljno usmerjeni spletni oglasi na podlagi napovedi klikov

Poslovni modeli, ki uporabljajo napovedi klikov in ciljno usmerjene oglase, pogosto sledijo politiki „klikni in kupi“. **Podjetja** kupujejo oglasni prostor na medijskih platformah in optimizirajo prikaz oglasov z analizo interesov in preferenc uporabnikov spletnih strani ter jim prikazujejo oglase, ki so zanje zanimivi. Namen je povečati relevantnost prikazanih oglasov na podlagi boljšega usklajevanja z interesi tistih, ki jih vidijo.

V tem primeru podjetje dobi plačilo le, če ljudje kliknejo na oglas in nekaj kupijo. Poleg tega podjetje uporablja umetno inteligenco za odkrivanje neprimernih vsebin v oglasih, kot so npr. oglasi za alkohol, strelno orožje ali politične vsebine.

Na področju računalniškega oglaševanja podjetje uporablja vrsto tehnik strojnega učenja. Za oceno verjetnosti, da bo uporabnik kliknil na oglas, prikazan v določenem kontekstu (optimizacija t. i. razmerja med prikazi in kliki), se interesi strank in ustreznost izdelkov merijo s kartiranjem zgodovine brskanja posameznikov in vzorcev transakcij. Poleg tega se informacije pridobivajo tudi na podlagi krmarjenja posameznikov po komercialnih spletnih straneh po vsem svetu, s katerimi oglaševalsko podjetje sodeluje. To se izvaja s pomočjo anonimiziranih piškotkov in sledilnikov tretjih oseb. Ti se nahajajo na spletnih straneh trgovcev in beležijo krmarjenje posameznikov po njih ter ustvarjajo sezname izdelkov, ki so jih videli in kupili.

Profili posameznikov so povezani z napravami, ki jih uporabljajo, čeprav so IP-naslovi anonimizirani. Ko je izdelek kupljen, algoritem sistema priporočevalca poskuša določiti druge izdelke, ki bi jih kupec prav tako lahko kupil. V tem primeru so „sveži“ podatki ovrednoteni višje kot starejši. Zgodovina brskanja se shranjuje za največ eno leto, saj se zanimanja spreminjajo in nakupi, starejši od enega leta, niso več nujno merodajni.

Oglasi, prikazani zadevni osebi, se takoj ustrezno prilagodijo in se razlikujejo glede na ustrezno spletno mesto, da se ujemajo tudi z vsebino slednjega. Ko je oglas objavljen, se nenehno analizira. Kombinacija elementov, ki se upoštevajo pri interesih posameznika, se potrdi ob nakupu. Podatki se izmenjujejo med različnimi platformami, kar vključuje medsebojno obveščanje po opravljenem nakupu, da se ustavi oglaševanje tega določenega izdelka. Če nakup ni opravljen, se formula pregleda in se algoritem dodatno prilagaja glede na sprotno spletno obnašanje posameznika.

Podjetje, ki je omenjeno v tem primeru, pričakuje, da si bo v prihodnosti znotraj svojega proračuna bolj prizadevalo za optimizacijo časovnega razporeda oglaševanja, kar bo izvedeno v točno določenem časovnem okviru. Obenem v podjetju pričakujejo, da se bodo bolj osredotočili na prikazane oglase, ki so imeli določen vpliv na potrošnike.

Drug primer temelji na evropski spletni tržnici, ki prek vrste specializiranih produktov povezuje kupce in prodajalce. V tem primeru se umetna inteligenca uporablja za optimizacijo oglaševalskih kampanj, za kategorizacijo izdelkov na podlagi oglasov, prikazanih na spletnem mestu tržnice, za izboljšanje izkušenj z iskalniki prek napovedi dopolnilnih in nadomestnih izdelkov ter za odkrivanje poskusov goljufije.

Podjetje uporablja strojno učenje za napovedovanje vrednosti klikov strank za nakup oglasnega prostora, ki je na voljo na dražbah v realnem času. Na podlagi teh primerov družba navaja, da ji umetna inteligenca omogoča sprejemanje odločitev, ki brez nje ne bi bile mogoče ali bi bilo treba znatno zmanjšati njihov pomen.

Ciljno usmerjeno komuniciranje s strankami in naročniki

V primeru maloprodajnega **podjetja**, ki se osredotoča na specializirano opremo, ki se prodaja v fizičnih trgovinah in na spletu, se neposredno trženje ali personalizirano oglaševanje uporablja za povečanje privlačnosti za stranke in hkrati za merjenje učinkovitosti določene oblike trženja ali oglaševanja.

Po podatkih tega podjetja oglasna e-poštna sporočila odpre v povprečju 20–30 % prejemnikov, zlasti ko ti ugotovijo, da se jim ponujajo primerni ali njim najljubši izdelki. Oglasna e-poštna sporočila se pošiljajo približno 250 000 registriranim posameznikom, sistem pa je namenjen ugotavljanju, kaj vsak izmed teh posameznikov šteje kot relevantno. To se izvede z analizo nakupov, ki so jih posamezniki opravili v preteklih šestih mesecih. 80 % prikazanih ponudb neposredno temelji na prejšnjih nakupih, medtem ko je 20 % novih predlogov, tj. alternativnih izdelkov iz iste kategorije kot prejšnji nakupi.

Podoben pristop uporablja banka pri pošiljanju e-poštnih sporočil svojim strankam. Sporočila s ponudbo določenih storitev ali izdelkov se pošiljajo samo določenim strankam. Podatkovni analitiki izračunajo verjetnost, da se stranka zanima za storitev ali izdelek. Če je ta verjetnost nad določenim pragom, bo stranka prejela sporočilo. Uporabljeni sistem še ne vključuje modelov strojnega učenja in ni popolnoma avtomatiziran. Te točke bodo upoštevane pri nadaljnjem razvoju sistema.

Tretji primer obravnava maloprodajnega trgovca z živili, ki uporablja kartice zvestobe za povečanje interakcij s strankami in personalizacijo ponudb. Sistemi kartic zvestobe lahko predvidijo, koliko strank bo verjetno izkoristilo ponudbo izdelkov. Sistem, ki je zajet v tem primeru, strankam predlaga tudi nove izdelke in sledi rezultatom teh predlogov. Kupce s podobnimi vedenjskimi vzorci razvrsti v segmente z namenom bolj prilagojenih predlogov.

Vsak teden lastniki kartic zvestobe podjetja po elektronski pošti, prek spletnega mesta ali mobilne aplikacije prejmejo prilagojene ponudbe, do ponudb pa lahko dostopajo tudi prek posebnih terminalov v trgovini. Sistem umetne inteligence na podlagi individualne zgodovine nakupov izbere ponudbe in priporoča nove artikle, ki bi lahko bili za kupca zanimivi in ga spodbudili k nakupu.


Male
Age 21
Male
Age 29
Female
Age 35
Female
Age 23
New Available

Končne opombe

1 Glej npr. Samoili, S. et al. (2020), ***AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*** (Opazovalnik umetne inteligence: opredelitev umetne inteligence: na poti k operativni opredelitvi in taksonomiji umetne inteligence), Luxembourg; Karanasiou, A. in Pinotsis, D. (2017), „**A study into the layers of automated decision-making: emergent normative and legal aspects of deep learning**“ (Poglobljena študija avtomatiziranega odločanja: nastajajoči normativni in pravni vidiki globokega učenja), *International Review of Law, Computers & Technology*, 2017, str. 170–187.

2 Glej FRA (2019), ***Facial recognition technology: fundamental rights considerations in the context of law enforcement*** (Tehnologija za prepoznavanje obraza: temeljni vidiki pravic v okviru kazenskega pregona), Luxembourg, Urad za publikacije, str. 9 in 22.

3 FRA (2019), ***Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*** (Kakovost podatkov in umetna inteligenca – blažitev pristranskosti in napak pri varstvu temeljnih pravic), Luxembourg, Urad za publikacije, junij 2019.

4 Svet OZN za človekove pravice (2019), Poročilo Posebnega poročevalca za izredno revščino in človekove pravice, Philip Alston, A/74/48037.

5 Eubanks, V. (2018), *Automating Inequality. How high-tech tools profile, police, and punish the poor* (Avtomatizacija neenakosti. Kako visokotehnološka orodja profilirajo, nadzirajo in kaznujejo revne), St. Martin's Press.

6 Redden, J., Dencik, L. in Warne, H. (2020), ***Datafied children welfare services: unpacking politics, economics and power*** (Pretvarjanje storitev za socialno varstvo otrok v podatke: razkritje političnih, gospodarskih vidikov in oblasti), *Policy Studies*.

7 Panoptykon Foundation (2015), ***Profiling the unemployed in Poland: Social and political implications of algorithmic decision making*** (Profiliranje brezposelnosti na Poljskem: socialne in politične posledice algoritemskega odločanja); glej tudi Algorithm Watch (2019), ***Poland: Government to scrap controversial unemployment scoring system*** (Poljska: vlada umika sporni sistem za ocenjevanje brezposelnosti).

8 OECD, glosar statističnih izrazov – opredelitev socialnih prejemkov, obiskano 5. avgusta 2020.

9 Richardson, J. H., „**CHAPTER IV, SOCIAL INSURANCE**“, *Economic and Financial Aspects of Social Security* (POGLAVJE IV: SOCIALNO ZAVAROVANJE, Gospodarski in finančni vidiki socialnega varstva; University of Toronto Press, 1960). Pieters, *Social Security*.

10 Za pregled pristojnosti EU na tem področju in Uredbe (ES) št. 883/2004 glej Paju, J. (2017), *The European Union and Social Security Law* (Pravo socialne varnosti Evropske unije), Oxford, Hart Publishing, poglavje 2.

11 Bakke, E. (2018), „Predictive policing: The argument for public transparency“ (Napovedno policijsko delo: argument v prid javni preglednosti), *New York University Annual Survey of American Law*, zvezek 74, str. 139–140; Ferguson, A. G. (2017), „**Policing Predictive Policing**“ (Nadzor napovednega policijskega dela), *Washington University Law Review*, zvezek 94, str. 1146–1150. Na primer, samo ena od petih žensk, ki so bile žrtev nasilja, je najhujši incident prijavila policiji. Glej FRA (2014), ***Violence against women: an EU-wide survey. Main results report*** (Nasilje nad ženskami: vseevropska raziskava. Poročilo o glavnih rezultatih), Luxembourg, Urad za publikacije, str. 61.

12 Joh, E. E. (2015), ***The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*** (Nova diskretnost nadzora: avtomatizirani sum, masovni podatki in nadzor), UC Davis Legal Studies Research Paper No. 473, str. 18.

13 Braga, A. et al (2019), „**Hot spots policing of small geographic areas effects on crime**“ (Učinki nadzora žarišč na majhnih geografskih področjih na kriminal), *Campbell Systematic Reviews*, zvezek 15(3).

14 Joh, E. E. (2015), *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing* (Nova diskretnost nadzora: avtomatizirani sum, masovni podatki in nadzor), UC Davis Legal Studies Research Paper No. 473, str. 17–18. Dostopno na: **SSRN**.

15 Bakke, E. (2018), „Predictive policing: The argument for public transparency“ (Napovedno policijsko delo: argument v prid javni preglednosti), *New York University Annual Survey of American Law*, zvezek 74, str. 137–138.

16 Hardyns, W. in Rummens, A. (2017), „Predictive Policing as a New Tool for Law Enforcement? Recent Developments and Challenges“ (Napovedno policijsko delo kot novo orodje kazenskega pregona? Najnovejši razvoj in izzivi), *Eur J Crim Policy Res*, str. 3, DOI: 10.1007/s10610-017-9361-2.

17 Meijer, A. in Wessels, M. (2019), „Predictive Policing: Review of Benefits and Drawbacks“ (Napovedno policijsko delo: pregled koristi in pomanjkljivosti), *International Journal of Public Administration* 42:12, str. 1032, DOI: 10.1080/01900692.2019.1575664.

18 Komisija pravniške zbornice o uporabi algoritmov v sistemu pravosodja (2019), ***Algorithms in the criminal justice system*** (Algoritmi v sistemu kazenskega prava), str. 36.

19 Newbold, J. (b. d.), „*Predictive Policing*“, ***„Preventative Policing“ or „Intelligence Led Policing“. What is the future?*** (Napovedno policijsko delo, preventivni nadzor, nadzor na podlagi obveščevalnih podatkov. Kakšna je prihodnost?).

20 Izjava št. 21, priložena Sklepni listini Medvladne konference, ki je sprejela Lizbonsko pogodbo, podpisano 13. decembra 2007.

21 Direktiva (EU) 2016/680 Evropskega parlamenta in Sveta z dne 27. aprila 2016 o varstvu posameznikov pri obdelavi osebnih podatkov, ki jih pristojni organi obdelujejo za namene preprečevanja, preiskovanja, odkrivanja ali pregona kaznivih dejanj ali izvrševanja kazenskih sankcij, in o prostem pretoku takih podatkov ter o razveljavitvi Okvirnega sklepa Sveta 2008/977/PNZ, UL L 119, 4.5.2016, str. 89–131.

22 Evropska komisija (2019), standardni Eurobarometer 92, poročilo, Evropejci in umetna inteligenca, str. 10.

23 Evropski forum pacientov (b. d.), ***The new EU Regulation on the protection of personal data: what does it mean for patients? A guide for patients and patients' organisations*** (Nova uredba EU o varstvu osebnih podatkov: kaj pomeni za paciente? Vodnik za paciente in organizacije pacientov).

24 Priporočilo Komisije (EU) 2019/243 z dne 6. februarja 2019 o evropski obliki izmenjave elektronskih zdravstvenih zapisov, UL L 39, 11.2.2019, str. 18–27.

25 Digital Health Society, ***Exchange of electronic health records across the EU*** (Izmenjava elektronskih zdravstvenih zapisov po EU), 19. februar 2020.

26 Sporočilo Komisije Evropskemu parlamentu, Svetu, Evropskemu ekonomsko-socialnemu odboru in Odboru regij, **Evropska strategija za podatke**, COM(2020) 66 final, Bruselj, 19. februar 2020.

27 Sklep št. 1082/2013/EU Evropskega parlamenta in Sveta z dne 22. oktobra 2013 o resnih čezmejnih nevarnostih za zdravje in o razveljavitvi Odločbe št. 2119/98/ES, UL L 293, 5.11.2013, str. 1–15.

28 Glej spletno stran Komisije o nalezljivih boleznih.

29 OECD in Evropska unija (2018), ***Healthcare at a glance: Europe 2018*** (Pregled zdravstva: Evropa 2018), str. 192.

30 Ehrenstein, V., Kharrazi, H., Lehmann, H. in Overby Taylor, C. (2019), „Obtaining Data From Electronic Health Records“ (Pridobivanje podatkov iz elektronskih zdravstvenih zapisov), v: Gliklich, R. E., Leavy, M. B. in Dreyer, N. A. (ur.), ***Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide*** (Orodja in tehnologije za interoperabilnost evidenc, evidence za vrednotenje izida pri pacientih: priročnik za uporabnike), 3. izdaja, Dodatek 2.

31 Za trenutno in potencialno uporabo teh podatkov v zavarovalniški industriji glej npr. Spender, A., Bullen, C., Altmann-Richer, L., Cripps, J., Duffy, R., Falkous, C., Farrell, M., Horn, T., Wigzell, J. in Yeap, W. (2019), „**Wearables and the internet of things: considerations for the life and health insurance industry**“ (Nosljivi pripomočki in internet stvari: razmislek za sektor življenjskega in zdravstvenega zavarovanja), *British Actuarial Journal* 24:22, str. 1–31.

32 Glej **vizualizacijo Svetovne zdravstvene organizacije**.
33 Glej kratek pregled različnih sistemov elektronskih zdravstvenih zapisov v Evropi z zornega kota medicinskih sester v Health Europa (2019), *The world of cloud-based services: storing health data in the cloud* (Svet storitev v oblaku: shranjevanje zdravstvenih podatkov v oblaku).
34 College of Europe (2018), *Transformation Health and Care in the Digital Single Market. Synopsis report of the public consultation* (Preobrazba zdravja in oskrbe na digitalnem enotnem trgu: zbirno poročilo javnega posvetovanja).
35 Evropska komisija (2016), ***Study on Big Data in public health, telemedicine and healthcare*** (Študija o masovnih podatkih v javnem zdravju, telemedicini in zdravstvenem varstvu); Pastorino, R., De Vito, C., Migliara, G., Glocker, K., Binenbaum, I., Ricciardi, W. in Boccia S. (2019), „**Benefits and challenges of Big Data in healthcare: an overview of the European initiatives**“ (Koristi in izzivi masovnih podatkov v zdravstvu: pregled evropskih pobud), *European Journal of Public Health*, zvezek 29, priloga izdaje št. 3, str. 23–27.
36 Ministrstvo za zdravje, socialno varstvo in šport Nizozemske (2016), ***Digitalization in health care and benefits for patient safety: Literature and web reports (2015-2016)*** (Digitalizacija v zdravstvu in koristi za varnost pacientov: Literatura in spletna poročila (2015–2016)).
37 To je v skladu z več poročili različnih podjetij za kibernetsko varnost v različnih časovnih obdobjih. Glej npr. SC Magazine (2019), *Healthcare leads in cost of data breaches* (Zdravstveni sektor je vodilni pri stroških kršitev varnosti podatkov); Williams, S. (2020), ***New report reveals 'wall of shame' in health care data breaches*** (Novo poročilo razkriva „zid sramote“ pri stroških kršitev varnosti podatkov v zdravstvu); Lovell, T. (2019), ***Statistics reveal healthcare is the sector most affected by personal data breaches*** (Statistika razkriva, da je zdravstveni sektor najbolj prizadet s kršitvami varnosti podatkov).
38 SC Magazine (2019), *Healthcare leads in cost of data breaches* (Zdravstveni sektor vodi pri stroških kršitev varnosti podatkov).
39 Sollie, A. (2016), ***Reuse and Sharing of Electronic Health Record Data with a focus on Primary Care and Disease Coding*** (Ponovna uporaba in souporaba podatkov iz elektronskih zdravstvenih zapisov s poudarkom na primarni oskrbi in kodiranju bolezni), doktorska disertacija na univerzi Vrije Univesiteit Amsterdam, str. 28–30.
40 Ehrenstein, V., Kharrazi, H., Lehmann, H. in Overby Taylor, C. (2019), „Obtaining Data From Electronic Health Records“ (Pridobivanje podatkov iz elektronskih zdravstvenih zapisov), v: Gliklich, R. E., Leavy, M. B., Dreyer, N. A. (ur.*),* ***Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User’s Guide*** (Orodja in tehnologije za interoperabilnost evidenc, evidence za vrednotenje izida pri pacientih: priročnik za uporabnike), 3. izdaja, Dodatek 2.
41 Househ, M., Aldosari, B., Alanazi Show, A., Kushniruk, A. in Borycki, E. M. (2017), *Big Data, Big Problems: A Healthcare Perspective* (Masovni podatki, veliki problemi: vidiki zdravstvenega varstva), *Studies in health technology and informatics 238*, str. 38.
42 Sartor, G. (2020), ***New aspects and challenges in consumer protection*** (Novi vidiki in izzivi varstva podatkov), študija za Odbor za notranji trg in varstvo potrošnikov, Tematski sektor za gospodarsko in znanstveno politiko ter kakovost življenja, Evropski parlament, Luxembourg.
43 Neudert, L. in Marchal, N. (2019), ***Polarisation and the use of technology in political campaigns and communication*** (Polarizacija in uporaba tehnologije v političnih kampanjah in komunikacijah), študija na zahtevo Odbora za prihodnost znanosti in tehnologije (STOA) pod vodstvom Oddelka za znanstvene napovedi znotraj Generalnega direktorata za parlamentarne raziskovalne storitve (EPRS) Sekretariata Evropskega parlamenta; Urad informacijskega pooblaščenca (ICO) (2018), ***Investigation into data analytics for political purposes*** (Preiskava uporabe podatkovne analitike v političnih kampanjah).
44 Svet Evrope (2019), **Izjava Odbora ministrov o manipulativnih zmogljivostih algoritemskih procesov**, Decl(13/02/2019)1.
45 Npr. Costello, R. Á. (2020), „**The Impacts of AdTech on Privacy Rights and the Rule of Law**“ (Učinki oglaševalske tehnologije na pravice do zasebnosti in vladavino prava), *Technology and Regulation*, 11–23; ENVP (2018), **Mnenje 3/2018, Mnenje ENVP o spletni manipulaciji in osebnih podatkih**.
46 Sartor, G. (2020), ***New aspects and challenges in consumer protection*** (Novi vidiki in izzivi varstva podatkov); Jabłonowska, A. et al. (2018), ***Consumer law and artificial intelligence. Challenges to the EU consumer law and policy stemming from the business’ use of artificial intelligence*** (Potrošniško pravo in umetna inteligenca: Izzivi potrošniškega prava in politik EU, ki izhajajo iz poslovne uporabe umetne inteligence), *EUI Working Papers, LAW 2018/11*.
47 Wachter, S. (2020), „Affinity Profiling and Discrimination by Association in Online Behavioural Advertising“ (Afinitativno profiliranje in diskriminacija zaradi pripadnosti v spletnem vedenjskem oglaševanju), *Berkeley Technology Law Journal*, zvezek 35, št. 2, 2020 (v pripravi), na voljo na **SSRN**.
48 Zuboff, S. (2018), *The Age of Surveillance Capitalism* (Doba kapitalizma nadzora), London; ENVP (2018), **Mnenje 3/2018, Mnenje ENVP o spletni manipulaciji in osebnih podatkih**.
49 Martin, G. (2011), „The importance of marketing segmentation“ (Pomen segmentacije trženja), *American Journal of Business Education*, zvezek 4, št. 6.
50 Lambe, K. in Ricks, B. (2020), ***The basics on microtargeting and political ads on Facebook*** (Osnove mikrociljanja in političnega oglaševanja na Facebooku).
51 Glej na primer informacije o **delavnici RecSys2020 na temo REVEAL 2020: Bandit and Reinforcement Learning from User Interactions** (Bandit in spodbujevalno učenje iz uporabniških interakcij) (dostop 7. avgusta 2020).
52 Direktiva 2006/114/ES Evropskega parlamenta in Sveta z dne 12. decembra 2006 o zavajajočem in primerjalnem oglaševanju, UL L 376, 27.12.2006, str. 21–27.
53 Evropska komisija, **Direktiva o zavajajočem in primerjalnem oglaševanju: Cilj direktive**.
54 Direktiva Evropskega parlamenta in Sveta 2005/29/ES z dne 11. maja 2005 o nepoštenih poslovnih praksah podjetij v razmerju do potrošnikov na notranjem trgu ter o spremembi Direktive Sveta 84/450/EGS, direktiv Evropskega parlamenta in Sveta 97/7/ES, 98/27/ES in 2002/65/ES ter Uredbe (ES) št. 2006/2004 Evropskega parlamenta in Sveta (Direktiva o nepoštenih poslovnih praksah), UL L 149, 11.6.2005, str. 22–39.
55 Direktiva 2006/123/ES Evropskega parlamenta in Sveta z dne 12. decembra 2006 o storitvah na notranjem trgu, UL L 376, 27.12.2006.
56 Glej spletno stran Evropske komisije za **sveženj o digitalnih storitvah**.

3.

PRAVNI OKVIR ZA VARSTVO TEMELJNIH PRAVIC NA PODROČJU UMETNE INTELIGENCE

Uporaba umetne inteligence, kot je predstavljena v štirih primerih uporabe, obravnavanih v **poglavju 2**, lahko vpliva na konkretne temeljne pravice (kot je natančneje opisano v **poglavju 4**). Popolna skladnost s temeljnimi pravicami je predpogoj za uporabo tehnologij umetne inteligence, ne glede na zadevno področje.

V tem poglavju je predstavljen splošni pravni okvir za varstvo temeljnih pravic v EU, ki ureja uporabo umetne inteligence, vključno z izbrano sekundarno zakonodajo EU in nacionalno zakonodajo (oddelek 3.1). Pravni okvir za varstvo temeljnih pravic zagotavlja normativno podlago in merila za oblikovanje, razvoj in uvajanje orodij umetne inteligence[1]. Pomaga pri določanju, ali je konkretna uporaba umetne inteligence v skladu s temeljnimi pravicami. Zahteve za upravičeno poseganje v temeljne pravice so opisane v **oddelku 3.3**.

3.1 PRAVNI OKVIR ZA VARSTVO TEMELJNIH PRAVIC, KI UREJA UPORABO UMETNE INTELIGENCE

Temeljni instrument pravnega okvira EU za varstvo temeljnih pravic, ki se nanaša na uporabo umetne inteligence, je Listina. Skupaj z nenapisanimi splošnimi načeli prava EU predstavlja glavni vir temeljnih pravic v Uniji. Listina vsebuje širok spekter temeljnih pravic in ima enako pravno veljavnost kot Pogodbi EU. Listina zavezuje vse institucije in organe EU, pa tudi države članice, ko izvajajo pravo Unije (člen 51(1) Listine)[2].

Številne pravice iz Listine so enake tistim iz Evropske konvencije o človekovih pravicah (EKČP)[3]. Njihova vsebina in obseg morata biti enaka ustreznim pravicam, zagotovljenim z EKČP (člen 52(3) Listine). Vendar ta določba ne preprečuje širšega varstva po pravu Unije.

Temeljne pravice je mogoče najti tudi v določbah Pogodb (glej npr. člen 6(2) PEU ter naslova V in X PDEU) in v sekundarnem pravu EU[4]. Te pravice so dodatno zavarovane z drugimi dokumenti sekundarne zakonodaje Unije.

Osrednji dokument sekundarnega prava EU v kontekstu umetne inteligence je splošna uredba o varstvu podatkov (GDPR – Uredba (EU) 2016/679)[5]. Ureja avtomatizirano obdelavo osebnih podatkov v Evropskem gospodarskem prostoru in obdelavo osebnih podatkov na kateri koli drug način, ki je del zbirke, v okviru prava EU. (Zato se GDPR ne uporablja za obdelavo podatkov, povezanih z nacionalno varnostjo.)

GDPR je povezana z direktivo o kazenskem pregonu, ki se uporablja za policijsko in pravosodno sodelovanje v kazenskih zadevah. Oba instrumenta EU vključujeta številne določbe o varstvu osebnih podatkov, ki določajo ključna načela obdelave podatkov, kot so zakonitost, pravičnost in preglednost[6].

Ali se uporablja zakonodaja EU na področju varstva podatkov, je odvisno od tega, ali poteka obdelava osebnih podatkov. Nekatere aplikacije umetne inteligence ne uporabljajo osebnih podatkov (primer so prometni podatki). Druge uporabljajo anonimizirane podatke. V teh primerih se zakoni o varstvu podatkov ne uporabljajo ali njihova uporaba ni popolnoma jasna[7]. Meja med osebnimi in neosebnimi podatki je zabrisana, ker obstaja določeno tveganje, da pride do ponovne identifikacije anonimiziranih podatkov, tj. da se anonimizacija izniči. Vendar pa je ponovna identifikacija praviloma nezakonita. Poleg tega morajo osebe, ki izvajajo ponovno identifikacijo podatkov, običajno vložiti veliko napora in morebiti potrebujejo dostop do dodatnih informacij o posameznikih, ki bi lahko bili vključeni v anonimiziran niz podatkov, za namene ponovne identifikacije. **Oddelek 4.2** podrobneje obravnava to temo v povezavi z rezultati razgovorov, ki so bili opravljeni za namene tega poročila.

Poleg pravnega reda EU o varstvu podatkov je za zaščito temeljnih pravic pri uporabi umetne inteligence in sorodnih tehnologij ključnega pomena evropska zakonodaja na področju varstva pred diskriminacijo. Člen 2 PEU določa, da je nediskriminacija ena od temeljnih vrednot Unije, člen 10 PDEU pa od Unije zahteva, da se bori proti diskriminaciji zaradi vrste razlogov. Poleg tega sta enakost pred zakonom in prepoved diskriminacije določena tudi v členih 20 in 21 Listine.

Za nameček obstaja tudi več direktiv EU, ki bolj specifično in podrobneje obravnavajo področje varstva pred diskriminacijo. Njihova področja uporabe se razlikujejo[8]. Mednje sodijo direktiva o enakosti pri zaposlovanju (2000/78/ES)[9], direktiva o rasni enakosti (2000/43/ES)[10], direktiva o enakem dostopu moških in žensk do blaga in storitev (2004/113/ES)[11] in prenovljena direktiva o enakosti spolov (2006/54/ES)[12].

Države članice EU so hkrati podpisnice drugih mednarodnih konvencij o človekovih pravicah (glej seznam konvencij v oddelku **Ključne ugotovitve in mnenja FRA**). Te vsebujejo pravno zavezujoče standarde in varovala, ki jih države morajo spoštovati, ko delujejo na področjih, ki niso v pristojnosti EU. Glavni tak instrument je EKČP, ki so jo ratificirale vse države članice EU. Spremljajo ga dodatni protokoli, ki jih je podpisala velika večina držav članic EU. EKČP ima širok doseg, ker se uporablja tudi za področja, ki jih pravo Unije ne ureja.

Dodaten vir vseevropskih obveznosti na področju varstva podatkov, ki zavezuje vse države članice EU, je Konvencija Sveta Evrope o varstvu posameznikov glede na avtomatsko obdelavo osebnih podatkov[13]. Ta je bila pred kratkim posodobljena[14].

Zaščitne ukrepe za varstvo temeljnih pravic prav tako vsebujeta sektorska zakonodaja EU in nacionalna zakonodaja. Pregled takšne bolj tehnične zakonodaje presega okvir tega poročila. Vendar je v tem poglavju navedenih nekaj primerov, ki so pomembni z vidika primerov uporabe, obravnavanih v tem poročilu. To dopolnjuje nekaj primerov nacionalnih zakonodaj petih zajetih držav članic EU.

Nobena od petih zajetih držav članic EU trenutno nima horizontalne zakonodaje, ki bi urejala konkretno področje umetne inteligence, čeprav države proučujejo morebitno potrebo po regulaciji. Nekatere države EU, kot je Finska, so izdale priporočila za samoregulacijo in razvoj standardov odgovornosti za zasebni sektor[15]. V Estoniji je bilo v posebni oceni ugotovljeno, da v bližnji prihodnosti ne bo potreben ločen zakon za področje umetne inteligence, saj sedanji pravni okvir zadostuje[16]. Vendar dolgoročna estonska strategija izpostavlja, da bo potrebno prilagajanje pravnega okolja, da se preprečijo nepotrebne ovire pri uvajanju umetne inteligence[17].

Situacija na področju sektorske zakonodaje, ki ureja uporabo umetne inteligence v različnih sektorjih, se med posameznimi državami članicami EU razlikuje. Vendar pa v zadnjem času na nacionalni ravni prihaja do aktivnega oblikovanja politik v zvezi z umetno inteligenco. Oblikovani so bili nacionalni akcijski načrti za področje umetne inteligence, ki ostajajo jedro razvoja politik v posameznih državah članicah. Nekatere države si prizadevajo za rast podjetništva[18]. Druge se osredotočajo na sprejemanje tržno usmerjenih politik, ki so združljive z agendo OZN za trajnostni razvoj do leta 2030[19]. Izobraževalne dejavnosti za spodbujanje umetne inteligence in povečanje javne uporabe umetne inteligence so pogosto opredeljene kot strateški cilji na področju umetne inteligence. Kot pomemben cilj so pogosto navedene tudi naložbe v raziskave in razvoj[20].

Čeprav domače razprave o možnih zakonodajnih reformah pozorno spremljajo evropske pobude, prihaja tudi do sprejemanja nacionalnih sektorskih zaščitnih ukrepov, ki so namenjeni varstvu temeljnih pravic. Na primer, Finska je namesto posameznih zakonov, ki bi urejali samo področje umetne inteligence, začela razmišljati o celoviti prenovi domačih zaščitnih ukrepov na področju človekovih pravic v javnem sektorju s predlogom širše in vsesplošne posodobitve zakonodaje.

V zvezi z obdelavo osebnih podatkov po zakonodaji o priseljevanju je finski odbor za ustavno pravo vložil predlog za okrepitev zaščitnih ukrepov finske ustave, ki bi med drugim odpravil ustavne pomanjkljivosti v zvezi s pravnim varstvom, odgovornostjo in nejasnostjo algoritmov pri avtomatiziranem odločanju. Kadar koli pride do avtomatizacije postopkov odločanja javnih organov, mora biti znotraj teh postopkov upoštevano ustavno načelo pravne države in hkrati ne sme biti ogroženo spoštovanje pravil o dobrem upravljanju in spoštovanju predpisanega postopka[21]. Ta predlog je formuliral vizijo zahtev, ki jih finska ustava predvideva za uporabo umetne inteligence in avtomatizirano odločanje znotraj javne uprave.

Raziskava je ugotovila tudi druge pobude in politike, povezane z umetno inteligenco in temeljnimi pravicami, v petih pregledanih državah članicah. Na primer, estonska listina o e-državi vsebuje povzetek pravic državljanov za boljšo elektronsko komunikacijo z agencijami. Naslavlja tudi umetno inteligenco v zvezi s pravico do seznanitve s tem, katere podatke javni organi zbirajo[22].

Podobno je tudi nizozemsko ministrstvo za notranje zadeve parlamentu predstavilo kratko politično poročilo o umetni inteligenci, javnih vrednotah in temeljnih pravicah[23]. Poročilo poudarja pristop, osredotočen na človeka, na področjih, kjer imajo aplikacije umetne inteligence močan vpliv na ljudi ali družbo kot celoto. Navaja tudi najpomembnejša tveganja umetne inteligence za temeljne pravice, kot je diskriminacija zaradi pristranskih podatkov ali omejitev medosebnih odnosov, če umetna inteligenca nadomesti nekatere oblike interakcij.

3.2 PRIMERI UPORABE

Socialno varstvo (primer uporabe št. 1)

Pri regulaciji na področju socialnega varstva so države članice EU poleg obstoječih horizontalnih uredb EU sprejele različna pravila za varstvo temeljnih pravic na tem področju (glej **oddelek 2.1**). Običajno gre za opredelitev pravil za obdelavo in varstvo osebnih podatkov za namene socialnih prejemkov in zavarovanja.

V Estoniji je bil na primer sprejet zakon o zavarovalništvu, ki se uporablja za vse vrste in oblike zavarovanja in ureja obdelavo in prenos osebnih podatkov v tem okviru. Določa, da lahko javni organi, ponudniki zdravstvenih storitev, zavarovalnice in tretje osebe posredujejo osebne podatke na zahtevo zavarovalnice, če so osebni zdravstveni ali sodni podatki potrebni za to, da zavarovalnica sklene zavarovalno pogodbo, ali če pravica in obveznost razkritja takih podatkov izhajata iz zakona. Področje uporabe tega zakona vključuje tudi prenos podatkov za namen obdelave znotraj sistemov umetne inteligence.

Zakon o socialnem varstvu vsebuje natančnejše določbe o varstvu podatkov oseb, ki potrebujejo socialno pomoč. Treba jih je obvestiti o obdelavi njihovih podatkov in pridobiti soglasje za nadaljnjo obdelavo. Vsaka oseba v dani ciljni skupini ima pravico, da zavrne obdelavo podatkov. Zakon o socialnem varstvu lokalnim organom omogoča tudi obdelavo (vključno z uporabo algoritmov) osebnih podatkov mladih, starih od 16 do 26 let, shranjenih v državnih registrih, z namenom ugotavljanja, kateri mladi niso zaposleni, se ne izobražujejo ali usposabljajo.

Na Finskem uporabo umetne inteligence v socialni oskrbi in zdravstvenem varstvu ureja zakon št. 552/2019 o sekundarni uporabi zdravstvenih in socialnih podatkov. Ta zakon temelji na standardih varstva in zaščite občutljivih osebnih podatkov, ki so določeni v GDPR. Njegov cilj je vzpostaviti pogoje za učinkovito in varno „obdelavo in dostop do osebnih zdravstvenih in socialnih podatkov za nekatere sekundarne namene, kot so raziskave in statistika, inovacije in razvoj, upravljanje znanja, poučevanje in načrtovanje s strani oblasti“[24]. Zakon ureja načine, na katere je registrirane zdravstvene podatke dovoljeno obdelovati, in načine, ki so nedopustni.

Za različne vrste socialnih prejemkov se uporablja več drugih zakonov. V Franciji se za obdelavo in dostop do osebnih podatkov, povezanih s socialnimi prejemki, uporablja Kodeks za odnose med javnostjo in upravo iz leta 2015 z manjšimi spremembami, ki so bile sprejete po začetku veljavnosti GDPR. Ta kodeks navaja, da morajo biti „algoritmi, ki jih uporabljajo javne uprave, objavljeni“ in da ima „oseba, za katero se uporablja avtomatizirano odločanje, pravico biti obveščena“[25].

Napovedno policijsko delo (primer uporabe št. 2)

V kontekstu napovednega policijskega dela so ključni zaščitni ukrepi z vidika temeljnih pravic vključeni v direktivo EU o kazenskem pregonu. Ta določa, na kakšen način morajo organi kazenskega pregona uporabljati nekatera glavna načela varstva podatkov, določena v GDPR[26]. Med njimi je zahteva, da upravljavci podatkov (tj. pristojni organi kazenskega pregona) posameznikom, na katere se nanašajo osebni podatki, zagotovijo informacije o dejavnostih upravljavca pri obdelavi podatkov, kot so identiteta in kontaktni podatki upravljavca podatkov, nameni obdelave in informacije o pravici do vložitve pritožbe (člen 13). V posebnih primerih morajo upravljavci podatkov zagotoviti tudi dodatne informacije, na primer pravno podlago za obdelavo, da lahko posamezniki, na katere se nanašajo osebni podatki, uveljavljajo svoje pravice. Pravica do dostopa (člen 14) upravljavcu podatkov nalaga, da na zahtevo posameznika, na katerega se nanašajo osebni podatki, potrdi, ali v zvezi z njim poteka obdelava podatkov. V tem primeru lahko posameznik, na katerega se nanašajo osebni podatki, dostopa do teh podatkov in zahteva tudi dodatne informacije, vključno z namenom in pravno podlago obdelave ter vrstami obdelanih osebnih podatkov. Pravica do obveščenosti in pravica do dostopa sta lahko v številnih primerih omejeni, med drugim za zagotavljanje preiskovanja, odkrivanja ali pregona kaznivih dejanj ali izvrševanja kazenskih sankcij, vključno z varovanjem pred grožnjami javni varnosti in njihovim preprečevanjem[27].

Poleg tega člen 11 direktive o kazenskem pregonu izrecno prepoveduje avtomatizirano sprejemanje odločitev[28]. Prepoved je podvržena omejitvam, če jih odobri zakonodaja EU ali nacionalna zakonodaja in so varovane pravice posameznika, na katerega se nanašajo osebni podatki, pri čemer mora biti vključena „vsaj pravic[a] do osebnega posredovanja s strani upravljavca" (za več podrobnosti glej **oddelek 4.2**).

V nekaterih primerih je področje uporabe nacionalne zakonodaje širše od področja uporabe direktive. Finski zakon o obdelavi osebnih podatkov v kazenskih zadevah in v povezavi z ohranjanjem nacionalne varnosti denimo krepi pravico do obveščenosti, saj ne razlikuje med informacijami, ki se zagotavljajo na splošno, in tistimi, ki se zagotavljajo v posebnih okoliščinah[29].

Zdravstveno varstvo (primer uporabe št. 3)

V zvezi z zaščitnimi ukrepi za varstvo temeljnih pravic na ravni EU pri uporabi umetne inteligence v zdravstvu GDPR pacientom daje pravico, da so obveščeni, deloma tako, da jim omogoča večji nadzor nad njihovimi osebnimi zdravstvenimi podatki. Takšni podatki se štejejo za „občutljive podatke"[30], ki jih najdemo denimo v zdravstveni dokumentaciji posameznikov[31]. Te pravice vključujejo pravico dostopa do lastnih osebnih (zdravstvenih) podatkov, ugovora obdelavi lastnih osebnih podatkov, popravka in izbrisa podatkov ter pravice v primeru kršitev[32].

V skladu z GDPR upravne globe za kršitve obdelave podatkov, vključno z zdravstvenimi podatki, niso dovoljene. Vendar na primer v Estoniji nacionalna zakonodaja omogoča najvišjo kazen v višini 400 000 EUR, če v takih primerih pride do uporabe prekrškovnega postopka. Podobne globe lahko v prekrškovnem postopku izreče tudi inšpektorat za varstvo podatkov[33].

V Franciji zakon o varstvu podatkov in zakon o javnem zdravju določata strožje zahteve od tistih, ki so določene v GDPR v zvezi z obdelavo zdravstvenih podatkov. Francoski zakon o varstvu podatkov je bil spremenjen z zakonom o posodobitvi zdravstvenega sistema, da se omogoči obdelava osebnih zdravstvenih podatkov za različne namene, če spadajo v področje uporabe ene od izjem od splošnega načela prepovedi obdelave občutljivih podatkov v skladu s členom 9 GDPR[34].

Ciljno usmerjeno oglaševanje (primer uporabe št. 4)

Pri obravnavi zaščitnih ukrepov za varstvo temeljnih pravic v zvezi s ciljno usmerjenim oglaševanjem in skritimi mehanizmi, zlasti v zvezi s profiliranjem, najpomembnejše določbe o temeljnih pravicah zagotavlja pravni okvir Unije za področje zasebnosti in varstva podatkov. Varstvo zasebnosti in osebnih podatkov ima status, ki ima prednost pred gospodarskimi koristmi. Zato so pravila o obdelavi (posebnih kategorij) osebnih podatkov pomembna za podjetja, ki delujejo na področju ciljno usmerjenega oglaševanja ali ga uporabljajo, saj podjetjem nalagajo določene obveznosti.

Glavni pravni določbi, ki vsebujeta pravila o varstvu osebnih podatkov v EU, sta GDPR in direktiva o zasebnosti in elektronskih komunikacijah (direktiva o e-zasebnosti), ki je *lex specialis* v razmerju do GDPR. GDPR se neposredno uporablja v vseh državah članicah EU, če ima podjetje sedež v EU in obdeluje osebne podatke ali ima sedež zunaj EU, vendar obdeluje podatke, ki se nanašajo na posameznike v Uniji.

Direktiva o e-zasebnosti zadeva obdelavo osebnih podatkov in varstvo zasebnosti v sektorju elektronskih komunikacij (npr. ko posamezniki uporabljajo računalnik, pametni telefon in tablični računalnik) in daje močan poudarek temeljnim pravicam. Evropska komisija je leta 2017 predlagala uredbo o e-zasebnosti, ki bi nadomestila sedanjo direktivo o e-zasebnosti[35]. Zakonodajni predlog bi razširil področje uporabe direktive in vključeval posebne določbe v zvezi z neželenim trženjem, piškotki in zaupnostjo.

3.3 ZAHTEVE ZA UPRAVIČENO POSEGANJE V TEMELJNE PRAVICE

V **poglavju 4** so na podlagi štirih primerov uporabe iz poglavja 2 izpostavljene izbrane temeljne pravice, ki jih zajema Listina in na katere še posebej vpliva umetna inteligenca. Pri večini teh pravic ne gre za absolutne pravice, zato zanje lahko veljajo omejitve v skladu s členom 52(1) Listine. V skladu s tem ta oddelek pred analizo, v kolikšni meri uporaba umetne inteligence vpliva na različne temeljne pravice, predstavlja splošne ukrepe, ki jih je treba izvesti za ugotavljanje, ali je pravica iz Listine lahko predmet omejitev ali ne.

Temeljne pravice, na katere vpliva umetna inteligenca, ki niso absolutne, so lahko predmet omejitev. Poseganje v take temeljne pravice je mogoče upravičiti le, če se spoštujejo zahteve iz Listine in EKČP v primeru tistih pravic iz Listine, ki ustrezajo pravicam, zajamčenim z EKČP (člen 52(3) Listine)[36].

V skladu s členom 52(1) Listine mora vsako omejevanje temeljnih pravic:

- biti predpisano z zakonom,
- dejansko ustrezati ciljem splošnega interesa, ki jih priznava Unija, ali biti potrebno zaradi zaščite pravic in svoboščin drugih,
- spoštovati bistveno vsebino pravice,
- biti potrebno in
- sorazmerno[37].

Sodišče EU je poleg tega poudarilo, da je treba pri vsaki omejitvi uresničevanja pravic in svoboščin, ki jih priznava Listina, upoštevati „bistvo“ teh pravic in svoboščin[38]. To pomeni, da je temeljne pravice mogoče do določene mere omejiti, vendar ne popolnoma ignorirati.

Ko se ugotovi, da ukrep ne krši bistvene vsebine pravice, je naslednji korak izvajanje testa nujnosti in sorazmernosti, opisanega v Listini, v zvezi z nebistvenimi vidiki te pravice[39]. Vsako poseganje v katero od pravic iz Listine je treba preučiti in preveriti, ali določenega legitimnega cilja ni mogoče doseči z drugimi sredstvi, ki manj intenzivno posegajo v zajamčeno pravico[40]. Podobne zahteve nalaga tudi EKČP, kakor jo razlaga Evropsko sodišče za človekove pravice (ESČP)[41]. Med njimi je koncept „bistva pravice", ki lahko izhaja iz cilja in namena EKČP kot celote[42]. V zvezi z uporabo novih tehnologij je ESČP v zadevi *S. in Marper proti Združenemu kraljestvu* ugotovilo, da bi morale države „zagotoviti pravo ravnovesje" med varstvom temeljnih pravic in razvojem novih tehnologij[43].

Glede na širok spekter načinov uporabe umetne inteligence v vsakdanjem življenju, kot je predstavljen v štirih izbranih primerih uporabe, bo morda treba oceniti širok spekter temeljnih pravic, ob upoštevanju različnih elementov, odvisno od konteksta in posebnega področja uporabe. Za oceno posledic za temeljne pravice so pomembni zlasti posebni namen, za katerega se uporablja umetna inteligenca, njena funkcionalnost, kompleksnost in obseg, v katerem se uporablja[44].

Končne opombe

1 Glej tudi van Veen, C. (2018), „Artificial Intelligence; What's Human Rights Got to Do with It?" (Umetna inteligenca: kako je povezana s človekovimi pravicami?), *Data & Society: Points* – blog raziskovalnega inštituta Data & Society, 14. maj 2018; Barfield, W. in Pagallo, U. (2020), *Advanced Introduction to Law and Artificial Intelligence* (Napredni uvod v pravo in umetno inteligenco), Cheltenham/Northampton, MA, Edward Elgar, 2020, str. 19–20.
2 Glej tudi sodbo Sodišča, Åklagaren proti Hansu Åkerbergu Franssonu [veliki senat], 26. februar 2013, točki 17, 20.
3 **Evropska konvencija o varstvu človekovih pravic in temeljnih svoboščin**, kakor je bila spremenjena s protokoloma št. 11 in št. 14, 4. november 1950, ETS 5.
4 Za pregled uporabe Listine glej FRA (2018a), ***Uporaba Listine Evropske unije o temeljnih pravicah v pravu in pri oblikovanju politike na nacionalni ravni***, Luxembourg, Urad za publikacije.
5 Uredba (EU) 2016/679 Evropskega parlamenta in Sveta z dne 27. aprila 2016 o varstvu posameznikov pri obdelavi osebnih podatkov in o prostem pretoku takih podatkov ter o razveljavitvi Direktive 95/46/ES (**Splošna uredba o varstvu podatkov**), UL L 119, 4.5.2016, str. 1–88.
6 Za več informacij glej FRA (2018), ***Priročnik o evropskem pravu varstva osebnih podatkov***. *Izdaja iz leta 2018*, Luxembourg, Urad za publikacije.
7 Glej npr. Hacker, P. (2020), *A Legal Framework for AI Training Data. Law, Innovation and Technology* (Pravni okvir za učne podatke umetne inteligence. Pravo, inovacije in tehnologija, v pripravi), na voljo na **SSRN**.
8 Za pregled evropskega protidiskriminacijskega prava glej FRA (2018), ***Priročnik o evropskem protidiskriminacijskem pravu***. *Izdaja iz leta 2018*, Luxembourg, Urad za publikacije.
9 Direktiva Sveta 2000/78/ES z dne 27. novembra 2000 o splošnih okvirih enakega obravnavanja pri zaposlovanju in delu, UL L 303, 2.12.2000, str. 16–22.
10 Direktiva Sveta 2000/43/ES z dne 29. junija 2000 o izvajanju načela enakega obravnavanja oseb ne glede na raso ali narodnost, UL L 180, 19.7.2000, str. 22–26.
11 Direktiva Sveta 2004/113/ES z dne 13. decembra 2004 o izvajanju načela enakega obravnavanja moških in žensk pri dostopu do blaga in storitev ter oskrbi z njimi, UL L 373, 21.12.2004, str. 37–43.
12 Direktiva 2006/54/ES Evropskega parlamenta in Sveta z dne 5. julija 2006 o uresničevanju načela enakih možnosti ter enakega obravnavanja moških in žensk pri zaposlovanju in poklicnem delu (preoblikovano), UL L 204, 26.7.2006, str. 23–26.
13 Konvencija o varstvu posameznikov glede na avtomatsko obdelavo osebnih podatkov, Strasbourg, 28. januar 1981 (ETS št. 108).
14 Protokol o spremembi Konvencije o varstvu posameznikov glede na avtomatsko obdelavo osebnih podatkov, Strasbourg, 10. oktober 2018 (CETS št. 223).
15 Delovna skupina za etiko organizacije AI Finland je pri etičnih izzivih dodatno poudarila pomen podjetij in samoregulacije. AI Finland, ***Etiikkahaaste* (Etični izziv), *Tekoäly on uusi sähkö*** (v finščini).
16 Republika Estonija (2019), ***Report of Estonia's AI Taskforce*** (Poročilo operativne skupine za umetno inteligenco Estonije), str. 38.
17 Estonska vlada **je začela s pripravo** dolgoročne strategije.
18 Glej npr. Nizozemska, Ministrstvo za gospodarske zadeve in podnebno politiko (2019), Strateški akcijski načrt za umetno inteligenco (*StrategischActieplan AI* – SAPAI).
19 Kot primer prizadevanj za prilagoditev ciljev razvoju trajnostnega trga glej Španija, Ministrstvo za znanost, inovacije in univerze (2019), Nacionalna strategija za umetno inteligenco (v španščini).
20 Za celovitejši pregled glej Evropska komisija (2019), Nacionalne strategije za umetno inteligenco, ali opazovalno skupino za politiko umetne inteligence OECD.
21 Finski odbor za ustavno pravo (2019), Mnenje Odbora PeVL 7/2019 Vp – HE 18/2019 vp: Osnutek predloga Parlamentu za zakon o obdelavi osebnih podatkov v upravi za priseljevanje in za povezane zakone.
22 Estonija, Nacionalni urad za revizijo in Pravosodni kancler (2018), *Everyone's Rights in e-State: The e-State Charter* (Pravice vsakogar v e-državi: Listina o e-državi).
23 Nizozemska, Ministrstvo za notranje zadeve in odnose kraljevine (2019), ***AI, publieke waarden en mensenrechten*** (Umetna inteligenca, javne vrednote in temeljne pravice; v nizozemščini).
24 Saxlin-Hautamäki, E. in Lilja, J. (2019), ***Secondary use of health data – the new Finnish Act*** (Sekundarna uporaba zdravstvenih podatkov – novi finski zakon).
25 de Donno, M. (2017), „The French Code *"Des Relations Entre Le Public Et L'Administration"*. A New European Era For Administrative Procedure?" (Francoski zakonik *Des relations entre le public et l'administration*. Nova evropska doba za upravni postopek?), *Italian Journal of Public Law* 2, str. 220–260.
26 Glej FRA (2018), **Preprečevanje nezakonitega profiliranja danes in v prihodnosti: priročnik**, Luxembourg, Urad za publikacije, preglednici 2 in 4.
27 Sajfert, J. in Quintel, T. (2017), ***Data Protection Directive (EU) 2016/680 for Police and Criminal Justice Authorities*** (Direktiva (EU) 2016/680 o varstvu podatkov za policijo in organe kazenskega pregona), na voljo na **SSRN**.
28 Upoštevajte, da se zdi, da se člen 11 direktive o kazenskem pregonu uporablja izključno za avtomatizirane odločitve, sprejete z avtomatizirano obdelavo. To pomeni, da ta varnostni ukrep ne velja, če je vključeno človeško posredovanje. Lynskey, O. (2019), ***Criminal justice profiling and EU data protection law: Precarious protection from predictive policing*** (Profiliranje v kazenskem pravu in zakonodaja EU o varstvu podatkov: negotovost pri varstvu pred napovednim policijskim delom), str. 21.
29 Angleški prevod je na voljo na **spletnem mestu Finlex**.
30 GDPR, uvodna izjava 10 in člen 9(1).
31 Evropski forum pacientov (b. d.), ***The new EU Regulation on the protection of personal data: what does it mean for patients? A guide for patients and patients' organisations*** (Nova uredba EU o varstvu osebnih podatkov: kaj pomeni za paciente? Vodnik za paciente in organizacije pacientov).
32 GDPR, čl. 15–17, 20–21 in 34.
33 White & Case (2019), ***GDPR Guide to National Implementation: Estonia*** (Priročnik za nacionalno izvajanje GDPR: Estonija).
34 Griguer, M. (2019), ***Processing health data in France: What to look out for after GDPR?*** (Obdelava zdravstvenih podatkov v Franciji: na kaj je treba paziti po sprejemu GDPR?)
35 Evropska komisija, **Predlog Uredbe Evropskega parlamenta in Sveta o spoštovanju zasebnega življenja in varstvu osebnih podatkov na področju elektronskih komunikacij ter razveljavitvi Direktive 2002/58/ES (uredba o zasebnosti in elektronskih komunikacijah)**, COM(2017) 10 final, Bruselj, 10.1.2017.
36 Listina, člen 52(3): „Kolikor ta listina vsebuje pravice, ki ustrezajo pravicam, zagotovljenim z Evropsko konvencijo o varstvu človekovih pravic in temeljnih svoboščin, sta vsebina in obseg teh pravic enaka kot vsebina in obseg pravic, ki ju določa navedena konvencija."

37 Kot je ponovno poudarilo in pojasnilo tudi Sodišče EU. Glej na primer sodbe z dne 16. decembra 2008, Satakunnan Markkinapörssi in Satamedia, C-73/07, točka 56; z dne 9. novembra 2010, Volker und Markus Schecke in Hartmut Eifert, združeni zadevi C-92/09 in C-93/09, točka 77; z dne 8. aprila 2014, Digital Rights Ireland Ltd proti Minister for Communications, Marine and Natural Resources in drugim in Kärntner Landesregierung in drugim, združeni zadevi C-293/12 and C-594/12, točka 52; z dne 6. oktobra 2015, Maximillian Schrems proti Data Protection Commissioner, C-362/14, točka 92, in z dne 17. decembra 2015, WebMindLicenses Kft. proti Nemzeti Adó-es Vámhivatal Kiemelt Adó- és Vám Főigazgatóság, C-419/14, točke 69 in 80–82.

38 Glej sodbo Sodišča z dne 6. oktobra 2015, Maximillian Schrems proti Data Protection Commissioner, C-362/14, točki 94–95, ki se sklicuje na člen 52(3) Listine. Glej tudi Scheinin, M. in Sorell, T. (2015), ***SURVEILLE Deliverable D4.10 – Synthesis report from WP4, merging the ethics and law analysis and discussing their outcomes*** (SURVEILLE, rezultat D4.10 – zbirno poročilo v okviru delovnega paketa 4, združitev analize etike in prava ter razprava o rezultatih), 7. april 2015, str. 9.

39 Glej npr. Brkan, M. (2019), „**The Essence of the Fundamental Rights to Privacy and Data Protection: Finding the Way Through the Maze of the CJEU's Constitutional Reasoning**" (Bistvo temeljnih pravic do zasebnosti in varstva podatkov: iskanje poti skozi labirint ustavnih obrazložitev Sodišča EU), *German Law Journal* 20 (2019), str. 867; Lenaerts, K. (2019), „**Limits on Limitations: The Essence of Fundamental Rights in the EU**" (Omejitve omejitev: bistvo temeljnih pravic v EU), *German Law Journal* 20 (2019), str. 779–794.

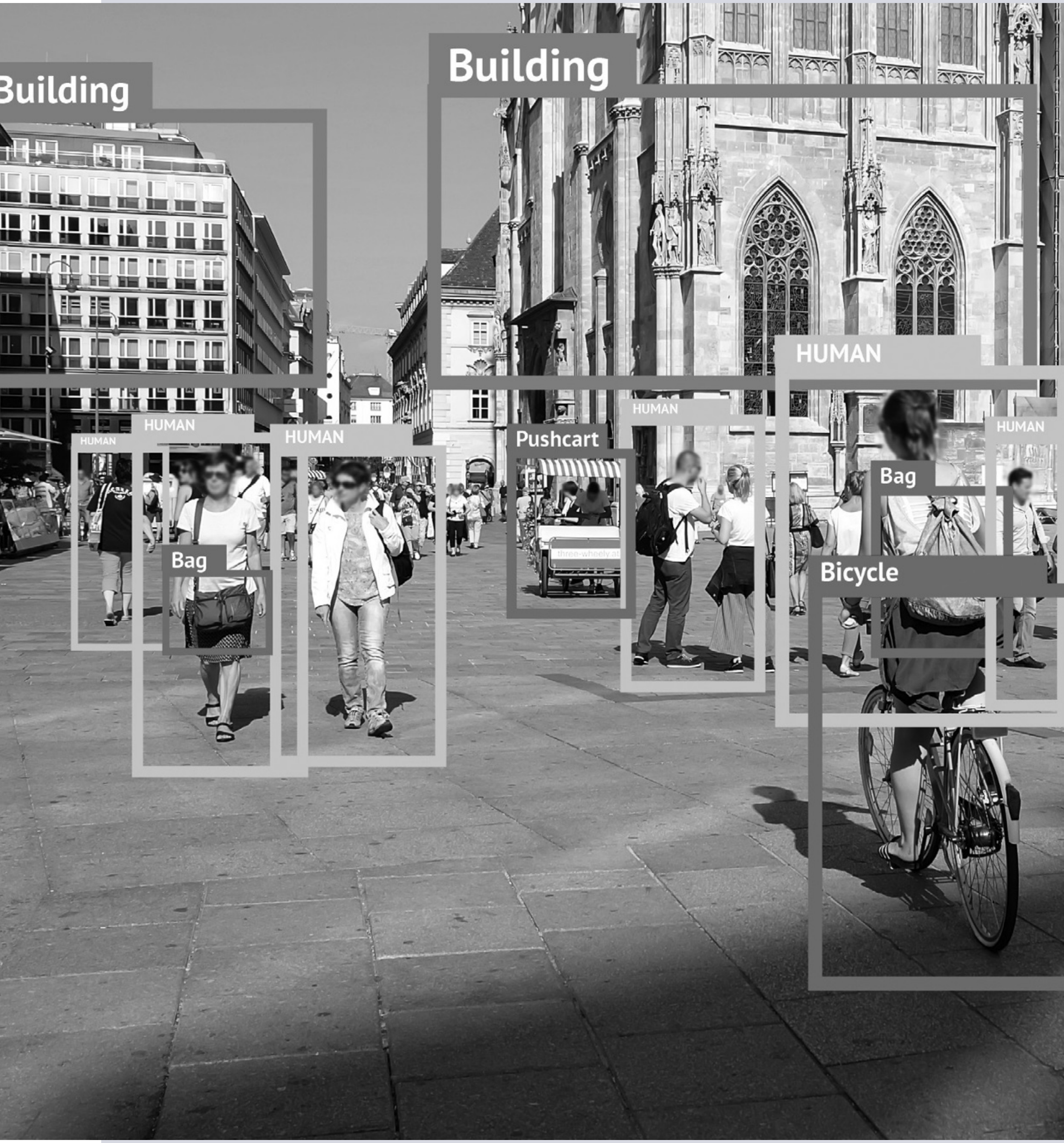
40 Sodišče EU, sodba z dne 8. aprila 2014, Digital Rights Ireland Ltd proti Minister for Communications, Marine and Natural Resources in drugim in Kärntner Landesregierung in drugim, združeni zadevi C-293/12 in C-594/12.

41 Glej npr. Khelili proti Švici, št. 16188/07, 18. oktober 2011; ESČP, S. in Marper proti Združenemu kraljestvu [veliki senat], št. 30562/04 in 30566/04, 4. december 2008; ESČP, K & T proti Finski, št. 25702/94, 12. julij 2001; ESČP, Z proti Finski, št. 22009/93, 25. februar 1997; ESČP, Huvig proti Franciji, št. 11105/84, 24. april 1990; ESČP, Leander proti Švedski, št. 9248/81, 26. marec 1987.

42 Scheinin, M. in Sorell, T. (2015), *SURVEILLE Deliverable D4.10 – Synthesis report from WP4, merging the ethics and law analysis and discussing their outcomes* (SURVEILLE, rezultat D4.10 – zbirno poročilo v okviru delovnega 4, združitev analize etike in prava ter razprava o rezultatih), 7. april 2015, str. 9.

43 ESČP, S. in Marper proti Združenemu kraljestvu [veliki senat], št. 30562/04 in 30566/04, 4. december 2008, točka 112.

44 Glej tudi ***Priporočilo Sveta Evrope CM/Rec(2020)1 Odbora ministrov državam članicam o posledicah algoritemskih sistemov za človekove pravice***, Dodatek, odst. A.8.


Building
Building
HUMAN
HUMAN
HUMAN
HUMAN
Pushcart
HUMAN
HUMAN
Bag
Bag
Bicycle

4.

VPLIV SEDANJE UPORABE UMETNE INTELIGENCE NA IZBRANE TEMELJNE PRAVICE

Uvajanje sistemov umetne inteligence zadeva širok spekter temeljnih pravic. Kot je razvidno iz **poglavja 2**, primeri uporabe, predstavljeni v tem poročilu, vključujejo vrsto tehnologij z različno stopnjo kompleksnosti in avtomatizacije. So v različnih fazah razvoja in se uporabljajo v različnih kontekstih, za različne namene in v različnem obsegu. Medtem ko so prizadete pravice odvisne zlasti od omenjenih dejavnikov, se na drugi strani pojavljajo številna horizontalna in sektorska temeljna vprašanja.

Poglavje se začne s splošnim pregledom tveganj, ki so jih prepoznali anketiranci, in njihove splošne ozaveščenosti o posledicah uporabe umetne inteligence za temeljne pravice. V poglavju so nato poudarjene izbrane temeljne pravice, na katere vplivajo tehnologije, povezane z umetno inteligenco, ob sklicevanju na štiri analizirane primere uporabe.

Analiza upošteva in predstavlja stališča, prakse in ozaveščenost o teh vprašanjih, izražene v razgovorih, ki so bili opravljeni v okviru priprave tega poročila. Anketiranci so bili najprej vprašani o splošnih tveganjih, ki jih vidijo pri uporabi umetne inteligence. Nadaljnja vprašanja so se nanašala na splošno ozaveščenost o temeljnih pravicah pri uporabi umetne inteligence in o konkretnejših posledicah za temeljne pravice, ki so bile večinoma povezane z varstvom podatkov, varstvom pred diskriminacijo in razpoložljivostjo pritožbenih mehanizmov.

4.1 ZAZNANA TVEGANJA

Pomembno je prepoznati, da se številna vprašanja dotikajo različnih pravic. Na primer, morebitna pristranska odločitev algoritma bi lahko vplivala na pravico do nediskriminacije, varstva osebnih podatkov in pravico do učinkovitega pravnega sredstva. Podobno je mogoče na določeno vprašanje gledati z vidika različnih pravic. Odločitev, sprejeta s pomočjo algoritma, zahteva podrobno razlago tako na podlagi pravice do varstva osebnih podatkov kot tudi do pravice do dobrega upravljanja ter pravice do učinkovitega pravnega sredstva in poštenega sojenja.

V zvezi z vprašanjem o splošnih tveganjih pri uporabi umetne inteligence anketiranci kot glavna tveganja niso vedno izpostavljali temeljnih pravic, čeprav so nekateri izpostavili povezane teme. Predstavniki zasebnega sektorja so kot glavno tveganje uporabe umetne inteligence najpogosteje omenjali netočnost, čemur sta sledili morebitna pristranskost in ustrezna pravna podlaga za obdelavo osebnih podatkov. Eden od anketirancev iz mednarodnega maloprodajnega podjetja je kot poslovno tveganje navedel dejstvo, da so evropski potrošniki zelo dobro seznanjeni s svojimi pravicami; ljudje se namreč ne obotavljajo vprašati o shranjevanju podatkov in avtomatiziranem odločanju. Če stranke niso ustrezno obveščene, se lahko pritožijo in podjetje lahko izgubi stranko. Poleg tega je sogovornik povedal, da predstavljajo kršitev zakona in morebitne globe, povezane s kršitvijo, dodatno poslovno tveganje.

Pri javni upravi je bila kot tveganje najpogosteje izpostavljena pristranskost, povezana z uporabo umetne inteligence. Poleg tega so pri tveganjih, povezanih z uporabo umetne inteligence, javni organi pogosto omenjali tudi netočnost in možnost ponovne identifikacije podatkov. Anketiranci, ki so uporabljali algoritme pri socialnih prejemkih, so navedli, da splošno tveganje predstavljajo nepravilni rezultati. Do tega lahko pride zaradi primerov, ki jih algoritem ne prepozna dobro, ali zaradi napak v vhodnih podatkih. Poudarili so tudi težave, povezane s prehodom s testiranja na uvajanje sistema, vključno s tehničnimi izzivi, potrebnimi viri in potencialnimi različnimi rezultati pri uvajanju.

Anketiranci, ki se ukvarjajo s ciljno usmerjenim oglaševanjem, so izpostavili tudi poslovna tveganja – na primer pri ponujanju neustreznih ali neprimernih vsebin. Eden izmed anketirancev je omenil morebitno izgubo nadzora nad avtomatiziranimi sistemi.

Poleg tega anketiranci navajajo izzive, povezane s težavami pri razlagi rezultatov iz sistemov umetne inteligence. Anketiranec iz svetovalnega sektorja se boji, da lahko tveganje, povezano s pomanjkanjem ali odsotnostjo zadostnega znanja in razumevanja umetne inteligence, povzroči ustavitev tekočih projektov zaradi nezmožnosti podjetij, da pojasnijo, kaj bodo algoritmi izvajali in s kakšnim namenom.

Drug anketiranec iz sektorja kazenskega pregona, ki preučuje morebitno uporabo umetne inteligence za podporo odločanju o vlogah za izdajo dovoljenj, pojasnjuje, da obstajajo tveganja glede tega, kako in zakaj sistem predlaga določen odziv. Na primer, anketiranec trdi, da je pri morebitni uporabi umetne inteligence za podporo odločanju o vlogah za izdajo dovoljenja za nabavo strelnega orožja pomembno ključno razumevanje razlogov za sprejem ne le negativnih odločitev, ampak tudi pozitivnih. Več razgovorov je pokazalo, da je glavna težava postavitev ustrezno usposobljenega osebja z zadostnim strokovnim znanjem za sledenje, razlago in interakcije s sistemom umetne inteligence.

To ugotovitev potrjujejo tudi rezultati raziskave Evropske komisije med podjetji v EU. V tej raziskavi je 85 % vprašanih kot oviro pri sprejemanju tehnologij umetne inteligence navedlo težave pri zaposlovanju novega osebja z ustreznim znanjem in spretnostmi, 80 % vprašanih kot oviro navaja tudi zapletenost algoritmov[1].

V zvezi z zmožnostjo pojasnjevanja odločitev, ki temeljijo na algoritmih, je anketiranec, ki je zaposlen v javni upravi, omenil, da pri sprejemanju odločitev ni mogoče pristati na manj kot popolno preglednost. Ne bi smelo biti prostora za dvom. V podobnem smislu anketiranec, ki dela v sektorju zasebnega zdravstva, omenja, da so na njihovem področju dela algoritmi, ki omogočajo samoučenje, prepovedani, ker je mogoče slediti le fiksnim algoritmom.

Druga tveganja, o katerih so anketiranci poročali, ne da bi pri tem navedli dodatne informacije, vključujejo kibernetsko varnost, kakovost podatkov, čezmerno spremljanje ljudi zaradi uporabe podatkov in algoritmov, izgubo delovnih mest zaradi avtomatizacije in profiliranje.

„Uporaba umetne inteligence lahko prinese številne koristi, vendar tudi tveganja, je kot jedrska energija.“
(anketiranec, zaposlen v zasebnem sektorju, Španija)

4.2 SPLOŠNA OZAVEŠČENOST O TEMELJNIH PRAVICAH IN PRAVNIH OKVIRIH V KONTEKSTU UMETNE INTELIGENCE

Vsi prebivalci EU se ne zavedajo svojih temeljnih pravic. Raziskava FRA o temeljnih pravicah kaže, da je za Listino slišal le nekaj več kot vsak drugi prebivalec EU (star 16 let ali več). Nekaj več ljudi, dva od treh, je slišalo za EKČP in Splošno deklaracijo o človekovih pravicah. To je morda zato, ker je EKČP starejša in bolj uveljavljena v splošnem zavedanju ljudi[2].

„[Naš način uporabe umetne inteligence] nikakor kakor koli ne vpliva [na človekove pravice]. V smislu postopka odločanja ni pomembno, ali odločitev sprejme stroj ali človek.“
(anketiranec, zaposlen v javni upravi, Estonija)

Večina vprašanih v okviru tega projekta priznava, da lahko uporaba umetne inteligence na splošno vpliva na temeljne pravice. Le redki so izjavili, da njihova uporaba umetne inteligence nima potencialnega vpliva na temeljne pravice oziroma da se takšnih posledic niso zavedali. Njihov odziv je odvisen od različnih načinov uporabe umetne inteligence, pa tudi od njihovega razumevanja temeljnih pravic.

Na primer, anketiranec, ki je delal na izračunih višine predvidenih pokojnin na podlagi strojnega učenja, pravi, da priprava statističnih podatkov nima vpliva na temeljne pravice, razen na področju varstva podatkov, kar je treba ustrezno obravnavati. Drug anketiranec, ki se ukvarja z algoritmi na področju socialnih prejemkov, trdi, da je učinek odvisen od „tega, kako široko so opredeljene človekove pravice“ – na primer pravica do prejemanja pravilne višine pokojnine.

„Ko zagotovimo spoštovanje vseh pravic v zvezi z varstvom podatkov, ne razumem, kakšno vlogo bi v tem pogledu lahko imele človekove pravice.“
(zasebno podjetje, Španija)

Nihče od anketirancev, ki delajo na področju ciljno usmerjenega oglaševanja, ne meni, da njihova uporaba umetne inteligence negativno vpliva na temeljne pravice. Eden od anketirancev, ki se je ukvarjal s ciljno usmerjeno komunikacijo s strankami, je izjavil, da je razlog za takšen odgovor pomanjkanje znanja o tem, kaj točno so temeljne pravice.

Praktično vsi anketiranci se zavedajo pravice do zasebnosti in varstva podatkov ter do nediskriminacije. Omenjene so bile tudi druge pravice, kot so človekovo dostojanstvo, pravica do poštenega sojenja in učinkovitega pravnega sredstva, čeprav le zelo na kratko.

„Teme nismo obravnavali, ker predpostavljamo, da ne gre za vprašanja človekovih pravic: vse dejavnosti potekajo znotraj pravnega okvira, vse dejavnosti so skladne z varstvom podatkov in dobrimi praksami, zato menimo, da vprašanja človekovih pravic niso povezana z uporabo teh sistemov.“
(javna uprava, Španija)

Podrobnejši pregled odgovorov anketirancev kaže na različne poglede različnih skupin anketirancev. Večina anketirancev, ki delajo v zasebnih podjetjih, omenja varstvo podatkov in varstvo pred diskriminacijo, vendar le redko omenja druge izzive v zvezi s pravicami. Anketirani iz podjetja, ki se ukvarja s ciljno usmerjenim oglaševanjem, povedo, da pozorno spremljajo vprašanja, povezana s svobodo govora in pravico do obveščenosti v smislu, da njihovo podjetje spodbuja te pravice. Po mnenju enega od anketirancev je razlog v tem, da objavljanje oglasov pomaga spletnim mestom z novicami in drugim spletnim mestom pridobiti finančna sredstva za nadaljevanje njihovega dela.

Obseg ozaveščenosti o pravicah je veliko širši med predstavniki javnega sektorja, ki se ukvarjajo z umetno inteligenco. Ti omenjajo druge pravice, kot sta človekovo dostojanstvo in domneva nedolžnosti.

Tisti, ki se ukvarjajo z uporabo sistemov umetne inteligence na različnih področjih, prav tako poudarjajo, da je uporaba sistemov med drugim zajeta tudi v nekaterih sektorskih zakonih. Sistem odločanja o dajatvah za primer brezposelnosti na primer ureja nacionalna zakonodaja o zavarovanju za primer brezposelnosti, upravne postopke in varstvo podatkov. Vendar pa nekateri anketiranci niso seznanjeni z nobenimi pravnimi standardi, ki bi veljali za njihovo uporabo umetne inteligence, ali o tem niso povsem prepričani.

V odsotnosti ureditve, specifične za področje umetne inteligence, več anketirancev omenja etične smernice in certifikacijske sheme. Nekateri se držijo obstoječih smernic in standardov, ki niso nujno posebej namenjeni urejanju umetne inteligence. To velja na primer za varnostni sistem informacijske tehnologije ISKE v Estoniji[3] ali za varnostni standard mednarodnih plačilnih sistemov (Payment Card Industry Data Security Standard) na področju finančnih storitev[4]. Anketiranci se sklicujejo tudi na standarde, ki so jih razvili Mednarodna organizacija za standardizacijo (ISO), Inštitut inženirjev elektrotehnike in elektronike (IEEE) ali Evropski odbor za standardizacijo (CEN).

Anketiranec, ki se ukvarja s ciljno usmerjenim oglaševanjem, trdi, da certificiranje na njegovem področju ni potrebno, ker objavljanje oglasov ni enako kot vprašanja, povezana z zdravstvenim sektorjem ali delom bank. Več anketirancev je izpostavilo, da njihove organizacije razvijajo (notranje) smernice.

„Mislim, da posebnih tehnologij, kot je umetna inteligenca, ni treba posebej urejati. Zadostujejo splošna načela in tehnološko nevtralna pravila.“
(zasebni sektor, Estonija)

Nekateri anketiranci omenjajo smernice, razvite na ravni EU in mednarodni ravni, kot so smernice strokovne skupine Evropske komisije na visoki ravni za umetno inteligenco, smernice OECD ali standardi Unesca. Nekateri se zavedajo razvoja, ki poteka na ravni EU in Sveta Evrope.

„Da, obstajajo kodeksi, in da, obstajajo postopki, vendar je oboje zastarelo, saj uporabljamo nekaj, kar smo ustvarili za analogni svet, v digitalnem svetu.“
(zasebni sektor, Španija)

Nekateri omenjajo tudi potrebo po posodobitvi sektorskih predpisov, ki je potrebna za uvajanje inovacij na področju umetne inteligence – na primer na področju zdravstva. Vendar pa en anketiranec odgovarja, da so obstoječi standardi zadostni in da umetne inteligence ni treba urejati ločeno.

4.3 ČLOVEKOVO DOSTOJANSTVO

Uporaba tehnologij umetne inteligence na splošno vključuje dolžnost spoštovanja človekovega dostojanstva, ki je osnova vseh temeljnih pravic, ki jih zagotavlja Listina[5]. Člen 1 Listine določa, da je človekovo dostojanstvo nedotakljivo in ga je treba spoštovati in varovati. Sodišče Evropske unije je v svoji sodni praksi potrdilo, da je temeljna pravica do dostojanstva del prava Unije[6].

Obdelavo osebnih podatkov, ki temelji na umetni inteligenci, je treba izvajati na način, ki spoštuje človekovo dostojanstvo. Na ta način postavimo človeka v središče vseh razprav in ukrepov, povezanih z umetno inteligenco. Namesto tehnologije mora biti v središču pozornosti človek, ki ustvarja novo tehnologijo in je pod njenim vplivom. Človekovo dostojanstvo kot izhodišče lahko pomaga zagotoviti, da bo uporaba umetne inteligence koristila vsem, na primer s podporo staranju in spodbujanju dostojanstvenega dostopa do zdravstvenega varstva.

Uporaba umetne inteligence lahko krši tudi druge tesno povezane pravice iz Listine, kot sta pravica do življenja (člen 2) in pravica do osebne celovitosti (člen 3). V zvezi s tem je pomembno razmisliti, kako preprečiti škodljivo uporabo umetne inteligence, da bi preprečili kršitve teh pravic, na primer kadar gre za uporabo umetne inteligence s strani ljudi, ki se ukvarjajo s kriminalnimi dejavnostmi, ali kadar se umetna inteligenca uporablja za orožje[7].

Poleg takšnih skrajnih primerov ohranjanje dostojanstva vključuje izogibanje izpostavljanju ljudi umetni inteligenci brez njihove vednosti in/ali informiranega soglasja, kar je tesno povezano z zasebnostjo in varstvom podatkov. Na primer, ko se o vlogah za socialne prejemke odloča z uporabo umetne inteligence, je treba ljudi o tem seznaniti (in pridobiti njihovo soglasje za uporabo umetne inteligence pri sprejemanju avtomatiziranih odločitev). Drug primer je, da se določen delež prebivalstva ob izpostavitvi biometričnim identifikacijskim sistemom počuti neprijetno. Zato bi lahko z uporabo takšnih sistemov, ne da bi omogočili zavrnitev sodelovanja, morda kršili njihovo dostojanstvo[8].

Le redki anketiranci iz javne uprave so pri obravnavi temeljnih pravic omenili pravico do dostojanstva. Eden od anketirancev je v zvezi z obravnavo morebitne uporabe umetne inteligence v zaporih navedel, da je treba v tem posebnem kontekstu najprej oceniti, ali bi bilo tveganje kršenja temeljnih pravic, kot je pravica do človekovega dostojanstva, preveliko. Drugi anketiranci so se na to pravico sklicevali le na splošno, ne da bi razmišljali o njenih posebnih vidikih v zvezi s konkretno uporabo umetne inteligence.

4.4 PRAVICA DO ZASEBNOSTI IN VARSTVA PODATKOV – IZBRANI IZZIVI

Pravica do spoštovanja zasebnega življenja in varstvo osebnih podatkov (člena 7 in 8 Listine) sta v središču razprav o temeljnih pravicah v zvezi z uporabo umetne inteligence. Čeprav sta tesno povezana, sta pravica do spoštovanja zasebnega življenja in varstvo osebnih podatkov ločena in predstavljata samostojni pravici. Govorimo o „klasični" pravici do varstva zasebnosti in „sodobnejši" pravici do varstva podatkov[9].

Obe sta namenjeni zaščiti podobnih vrednot, tj. avtonomije in človekovega dostojanstva posameznikov, tako da omogočata posamezniku osebno sfero, v kateri lahko svobodno razvija svojo osebnost, razmišlja in oblikuje svoje mnenje. Ti pravici predstavljata bistven predpogoj za uveljavljanje drugih temeljnih pravic, kot so svoboda misli, vesti in vere (člen 10 Listine), svoboda izražanja in obveščanja (člen 11 Listine) ter svoboda zbiranja in združevanja (člen 12 Listine)[10].

Glede na to, da ti pravici nista absolutni pravici, sta lahko omejeni. Vendar pa mora biti vsak poseg ustrezno utemeljen[11] in ne sme ogrožati bistvene in neodtujljive vsebine te pravice[12], kot je pojasnjeno v **oddelku 3.3**.

Pojem „zasebnega življenja" ali „zasebnosti" je zapleten in širok ter ne omogoča izčrpne opredelitve. Zajema telesno in duševno celovitost osebe in zato lahko vključuje več vidikov fizične in socialne identitete osebe[13]. Tudi področje interakcij posameznika z drugimi, celo v javnem kontekstu, lahko v nekaterih primerih spada v zasebno sfero. V drugih okoliščinah je ESČP uporabilo koncept „razumnega pričakovanja zasebnosti" – ki se nanaša na obseg, v katerem lahko ljudje pričakujejo zasebnost v javnih prostorih, ne da bi bili podvrženi nadzoru – kot eden od dejavnikov, čeprav ne nujno dokončen, za odločanje o kršitvi pravice do spoštovanja zasebnega življenja. Vendar se zdi, da sta njegova pomembnost in področje uporabe omejena[14]. Podobno po mnenju Odbora OZN za človekove pravice samo dejstvo, da so udeleženci v skupini na javnem mestu, ne pomeni, da njihove zasebnosti ni mogoče kršiti. Enako velja za primere, ko družbeni mediji zbirajo informacije o sodelovanju v miroljubnih zborovanjih[15].

Razširjena uporaba tehnologij umetne inteligence bi lahko v primeru njihovega nadaljnjega razvoja sprožila nerešena vprašanja in nove pomisleke glede pravice do spoštovanja zasebnega življenja. Tehnologije umetne inteligence lahko spremenijo naš pogled na zasebnost. Algoritemska orodja lahko predvidijo in razkrijejo informacije o vedenju ljudi na načine, ki so bili doslej nepredstavljivi, ne da bi se ljudje sploh zavedali, da dajejo takšne informacije. Osebni podatki, pridobljeni prek interneta, se lahko na primer uporabijo za ciljno usmerjeno oglaševanje, kar vzbuja veliko pomislekov glede temeljnih pravic[16]. Posebno zaskrbljenost vzbujajo vprašanja, povezana z izmenjavo osebnih podatkov prek aplikacij za pametne telefone, vključno z različnimi možnimi škodljivimi učinki, kot so manipulacija in izkoriščanje ranljivosti, diskriminacija, varnostna vprašanja in goljufije (npr. kraja identitete) ter manjše zaupanje v digitalno gospodarstvo[17].

Uporaba tehnologij umetne inteligence pogosto pomeni računalniško obdelavo velikih količin osebnih podatkov. Ta pa pomeni poseg v pravico do varstva osebnih podatkov iz člena 8 Listine (ki vključuje že obstoječo zakonodajo EU na področju varstva podatkov) ter pravico do zasebnega življenja v skladu s členom 7 Listine in členom 8 EKČP.

Ozaveščenost o vprašanjih varstva podatkov in uporaba osebnih podatkov

V EU je 69 % ljudi že slišalo za GDPR[18]. V nasprotju s tem pa so bili skoraj vsi anketiranci seznanjeni z GDPR in so razpravljali o vprašanjih varstva podatkov. Pravila o varstvu podatkov, ki izhajajo iz GDPR in nacionalne zakonodaje, so nedvomno najbolj znane in uveljavljene pravice na področju umetne inteligence. Druge temeljne pravice so nekoliko manj poznane.

Pri obravnavi pravnega okvira, ki ureja uporabo umetne inteligence, je večina anketirancev omenila le pravila o varstvu podatkov ter nekatere sektorske zakone. Nekateri so jasno povedali, da razen predpisov o varstvu podatkov ni nobenega drugega pravnega okvira. Anketiranec, ki dela za špansko javno upravo, pravi: „Zanašamo se na predpise in standarde s področja varstva podatkov, to je trenutno vse, kar je na voljo.“

Eden izmed anketirancev, ki je preučeval uporabo slikovnega diagnostičnega orodja, je izrazil stališče, da bi GDPR lahko prestavljala oviro pri nadaljnjih raziskavah. Anketiranec je navedel, da se v bolnišnici, ki uporablja orodje za podporo diagnozi po kapi, uporabljajo jasna pravila o varstvu podatkov, vendar ni bil prepričan, ali se zahteva potrjevanje varstva podatkov.

Drugi so se sklicevali na bolj splošne smernice za varstvo podatkov ali navedli, da s takšnimi dokumenti niso bili seznanjeni.

Vsi anketiranci, ki delajo na področju ciljno usmerjenega oglaševanja, se zavedajo vprašanj zasebnosti in varstva podatkov. Čeprav niso vsi odgovorni za področje varstva podatkov v svojih podjetjih, so vsi seznanjeni s prizadevanji za varstvo podatkov in zasebnosti. Eden od anketirancev je omenil, da se v nasprotju s prejšnjimi leti osebni podatki zdaj shranjujejo na veliko bolj varen način in se z njimi ravna bolj skrbno. Pozornost je bila namenjena pravilnemu ravnanju s privolitvijo za obdelavo podatkov. Posledično je prisotna visoka raven ozaveščenosti o varstvu podatkov in vprašanjih zasebnosti, povezanih z uporabo umetne inteligence.

Vendar pa se zakonodaja s področja varstva podatkov uporablja le v primerih obdelave osebnih podatkov. Uporaba anonimiziranih podatkov pri razvoju orodij umetne inteligence (tj. kot učnih podatkov) je v mnogih primerih najverjetneje dovoljena in ne pripelje do uporabe GDPR.

Raziskave kažejo, da se lahko podatki v številnih primerih tudi deanonimizirajo[19]. Vendar pa takšni poskusi pogosto zahtevajo strokovno znanje in morebitne

„V povezavi z izvajanjem GDPR smo bili nekoliko zaskrbljeni, vendar je to na koncu pomenilo upravljanje zbirk podatkov in pravic dostopa [...]. To je dobro opozorilo, da vsega ne moremo ali ne bi smeli narediti.“
(javna uprava, Finska)

„Pravzaprav me skrbi, da bi GDPR lahko predstavljala oviro pri raziskovanju umetne inteligence. Nekaterih velikih podatkovnih baz, ki smo jih prej uporabljali, žal pri naših raziskavah ne moremo več uporabljati.“
(zasebno podjetje, Nizozemska)

„Obstaja GDPR, vendar ne predpisuje posebnih pravil. Daje načela, a na koncu gre za etična vprašanja in razlago.“
(zasebno podjetje, Estonija)

dodatne informacije ter so nezakoniti. Čeprav nezakonitost deanonimizacije nujno ne izključuje uporabe GDPR, je pomembneje razmisliti, ali je ponovna identifikacija anonimiziranih podatkov razumno verjetna[20]. Anonimiziranje podatkov je le en vidik varstva zasebnosti posameznikov, na katere se nanašajo osebni podatki. Pri ocenjevanju tveganja ponovne identifikacije so pomembni tudi drugi vidiki, ki jih je treba upoštevati pri razširjanju anonimiziranih podatkov. Ti vključujejo podatke o uporabnikih podatkov, namenih uporabe ter ustvarjenih rezultatih[21].

V intervjujih anketiranci niso bili vedno povsem jasni glede uporabe osebnih podatkov. Pogosto so uporabljene podatke opisovali le površno, kot je navedeno v poglavju 2. V več primerih so anketiranci navedli, da uporabljajo neosebne podatke ali anonimizirane podatke, pri čemer so trdili, da varstvo podatkov v takih primerih ne pride v poštev. Poljavna organizacija, ki se ukvarja z okoljskim upravljanjem, na primer uporablja zbirne podatke o porabi vode za napovedi glede porabe vode, ki temeljijo na strojnem učenju. Ti podatki niso na voljo na individualni ravni.

Drugi anketiranci so povedali, da ne uporabljajo osebnih podatkov, čeprav ti prvotno izvirajo od posameznikov. Orodje, ki podpira izvajanje inšpekcij v restavracijah z zbiranjem podatkov iz spletnih virov, po navedbah sogovornika ne uporablja nobenih osebnih podatkov. Po drugi strani je povedal, da je pri spletnem rudarjenju podatkov treba biti previden, saj podatki na spletu kljub javni razpoložljivosti lahko vključujejo osebne podatke, kot so uporabniška imena.

„Odlično bi bilo, če bi lahko nekaj podatkov uspeli pridobiti iz druge storitve, saj jih strankam ne bi bilo treba ponovno vpisovati, ampak do kam sega meja pri možnostih ponovne uporabe podatkov?“

(javna uprava, Finska)

V drugem primeru zavarovalnica uporablja klepetalnega bota za učinkovitejši stik s strankami. Podatki, ki se uporabljajo za učenje sistema, so klepetalni protokoli (dnevniki pogovorov), ki niso povezani z osebnimi podatki. Vendar pa bo v tem primeru po mnenju anketiranca v prihodnosti morda mogoča povezava teh podatkov z osebnimi podatki.

Podjetja, ki se ukvarjajo s ciljno usmerjenim spletnim oglaševanjem, omenjajo uporabo (psevdo)anonimiziranih podatkov. To se izvede na primer z izključitvijo imen in oznak na področju socialne varnosti ter šifriranjem podatkov. Po navedbah enega od anketirancev identiteta potrošnikov za podjetje ni pomembna.

Medtem ko nekateri navajajo, da uporabljajo neosebne ali anonimizirane podatke, pri drugih to ni mogoče, ker se podatki uporabljajo za napovedovanje ali odločanje o konkretnih posameznikih. Na primer, anketiranec iz podjetja, ki se ukvarja z ocenjevanjem kreditne sposobnosti, je omenil, da mora pri svojih ocenah poznati identiteto potrošnikov. V tem primeru je to še pomembnejše kot pravica biti pozabljen, meni anketiranec.

Izčrpna razprava o vprašanjih varstva podatkov v tem poročilu ni mogoča. Vendar pa sta se med razgovori jasno pokazala dva vidika: avtomatizirano odločanje, povezano s pravico do človeškega nadzora, in pravica do pridobitve pomembnih informacij, ko so odločitve avtomatizirane.

Avtomatizirano odločanje

Člen 22 GDPR in člen 11 direktive o kazenskem pregonu na splošno prepovedujeta avtomatizirano odločanje, kar se nanaša na vsako „odločitev, ki temelji zgolj na avtomatizirani obdelavi, vključno z oblikovanjem profilov, ki ima pravne učinke v zvezi z določenim posameznikom ali na podoben način nanj znatno vpliva". V skladu s členom 22 GDPR je potrebna izrecna privolitev, kadar odločitve temeljijo zgolj na avtomatizirani obdelavi in imajo pravne ali podobne pomembne učinke na ljudi in če tako avtomatizirano odločanje ni dovoljeno z zakonom. Edini pogoj v skladu z direktivo o kazenskem pregonu je, da obdelavo dovoljuje pravo Unije ali nacionalno pravo (člen 11). Za odločitev, ki naj se ne šteje za popolnoma avtomatizirano, oba instrumenta zahtevata človeški nadzor s strani upravljavca[22].

Vendar je koncept „avtomatiziranega" odločanja težko oprijemljiv in zahteva nadaljnjo razpravo in raziskave. V nekaterih primerih je lahko na primer človeško posredovanje omejeno na „podpisovanje" rezultatov, ki jih ustvari sistem umetne inteligence, zaradi česar je odločanje praktično povsem avtomatizirano[23]. Pomembno je, da človeški nadzor ne sme biti sestavljen zgolj iz podpisovanja algoritemsko ustvarjenih priporočil ali rezultatov. Mora biti opravljen s strani nekoga, ki je „pooblaščen in pristojen za spreminjanje odločitev" ob upoštevanju vseh ustreznih podatkov, ki so na voljo[24]. Če ljudje pregledujejo in potencialno spreminjajo rezultate, pridobljene na podlagi sistema, mora biti celoten postopek tudi ovrednoten.

Raziskave kažejo, da ljudje rezultate algoritmov spreminjajo predvsem takrat, ko ti niso v skladu z njihovimi stereotipi[25]. Takšno vedenje ogroža možno dodano vrednost avtomatizirane obdelave, saj naj bi bila ta morebiti natančnejša ali celo pravičnejša od človeka. Prav tako lahko manjšinske skupine postavi v slabši položaj, zato je pomembno tudi v luči varstva pred diskriminacijo (kar obravnavamo v nadaljevanju).

Na splošno obstaja nekaj nesoglasij glede natančnega področja uporabe teh določb prava EU na področju varstva podatkov in glede tega, ali uvajajo splošno prepoved nekaterih vrst avtomatiziranih odločitev in ali posameznikom, na katere se nanašajo osebni podatki, zagotavljajo nekatere pravice v okviru določenih vrst odločanja, ki temeljijo na umetni inteligenci[26].

Uporaba algoritmov na področju socialnih prejemkov, zdravstva in napovednega policijskega dela zagotovo ima možne pravne ali druge pomembne posledice. Razgovori kažejo, da se tisti, ki delajo na teh področjih, dobro zavedajo koncepta človeškega nadzora pred sprejetjem odločitve s podporo umetne inteligence.

Številni anketiranci navajajo, da ne prihaja do sprejemanja avtomatiziranih odločitev. Ena od izjem je avtomatizacija odločitev o nadomestilih za primer brezposelnosti, ki so na podlagi nacionalne zakonodaje v primerih, ki ne vključujejo diskrecijske pravice, popolnoma avtomatizirane. V drugem primeru, ki se nanaša na drugo državo članico, so avtomatizirane samo pozitivne odločitve pri podeljevanju pomoči študentom, ki temeljijo na vnaprej določenih pravilih. V tem primeru vse negativne odločitve sprejemajo ljudje. Oba primera se nanašata na odločitve na podlagi pravil, ki ne vključujejo uporabe statističnih podatkov ali strojnega učenja.

Drug vprašani, ki se ukvarja s preizkušanjem uporabe sistemov umetne inteligence, vključno s strojnim učenjem, na področju socialnih prejemkov, omenja, da bi lahko takšne prakse negativno vplivale na enakost. To je zato, ker avtomatizacija naredi človeško vedenje vidno, vključno z obstoječimi pristranskimi praksami. Zaradi tega so potrebni previdnostni ukrepi in posledično organizacija v tem primeru dovoljuje samo odločitve, ki jih sprejme človek.

Anketiranci, ki delajo v zdravstvu, so prav tako izpostavili tveganja, povezana z avtomatizacijo odločitev. Anketiranec, ki preučuje uporabo orodja za podporo pri diagnozi možganske kapi, meni, da je pomembno, da se v izogib tveganju avtomatizacije ali potrditvene pristranskosti ne zanašamo na sistem. Opozarjajo, da lahko začetne pozitivne izkušnje z aplikacijo spodbudijo uporabnike, da se nanjo začnejo prehitro zanašati in lastni oceni slik na ta način posvetijo manj pozornosti. Podobne pomisleke so izrazili tudi drugi anketiranci. Eden od anketirancev, ki preverja možnost uporabe orodja za analizo slik z vidika verjetnosti za prisotnost določenih poškodb, ugotavlja, da tehnologija podpira diagnozo preprostih primerov, a da je zlasti v zapletenih primerih strokovno znanje zdravnikov zelo pomembno in zaupanja vredno.

Za ciljno usmerjeno oglaševanje se pogosto šteje, da naj ne bi imelo pomembnega vpliva na ljudi. Kljub temu do tega lahko pride, kot primer lahko navedemo situacije, v katerih se določene ranljivosti posameznika uporabljajo za uspešno oglaševanje. Upoštevanje ranljivosti je še posebej pomembno za ljudi iz prikrajšanih skupin, ki se morda ne zavedajo, da neposredno trženje lahko zavrnejo (glej okvir), ali ne poznajo svoje pravice, da lahko vplivajo na avtomatizirano sprejemanje odločitev.

Ker na tem področju ni sodne prakse, je potrebnih veliko več informacij in raziskav, da se ugotovi vpliv takšnih avtomatiziranih odločitev (tj. kateri oglas se bo prikazal ter komu, kdaj, kako in zakaj se bo prikazal). Odgovarjanje na ta vprašanja je zahtevno, saj ciljno usmerjeno oglaševanje temelji na zelo zapleteni tehnologiji in je zelo obsežno.

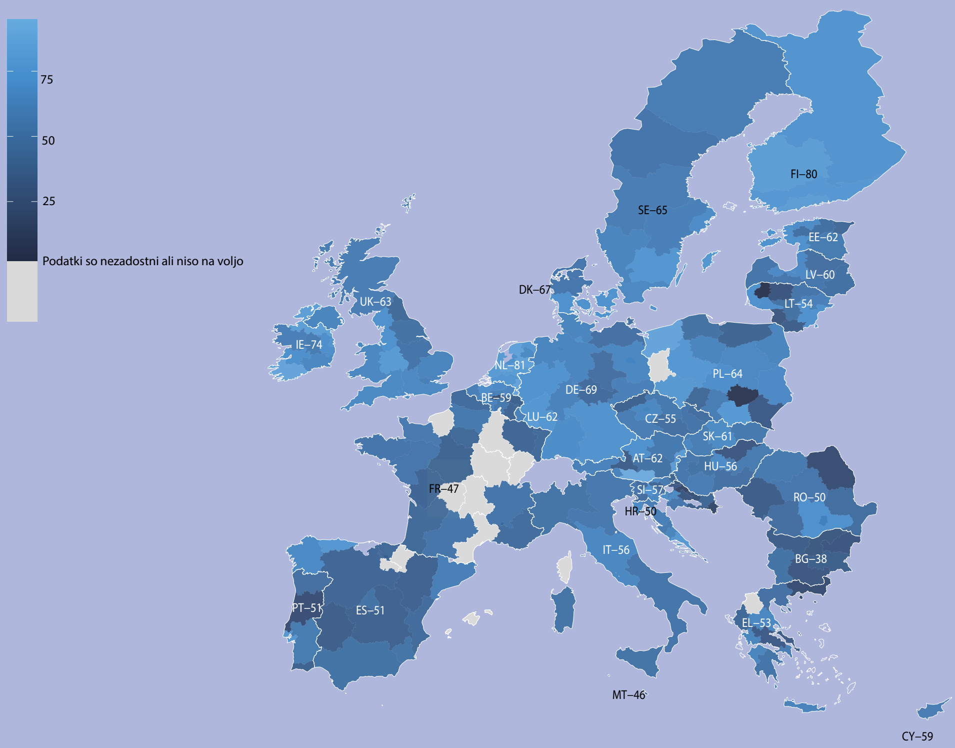
Ozaveščenost o pravici do zavrnitve neposrednega trženja med splošnim prebivalstvom

V okviru raziskave Eurobarometra so bili leta 2019 prebivalci EU vprašani, ali se zavedajo svoje pravice do zavrnitve neposrednega trženja. Na splošno je samo 59 % državljanov EU slišalo za to pravico (24 % jo je uveljavljalo). Vendar pa lahko ljudje svojo pravico uveljavijo le, če so o njej obveščeni – kar je še pomembnejše, če postane neposredno trženje zaradi uporabe strojnega učenja še toliko učinkovitejše.

Ravni ozaveščenosti se znotraj EU močno razlikujejo. Odstotek ljudi, ki se zavedajo svoje pravice do zavrnitve neposrednega trženja, sega od 38 % v Bolgariji do 81 % na Nizozemskem. Na sliki 4 so prikazani odstotki. Na podlagi analize podatkov Eurobarometra, ki jo je izvedla FRA, je prav tako izpostavljeno, da obstajajo velike razlike tudi znotraj posameznih držav, kadar te razdelimo na posamezne regije.

V nekaterih regijah je za omenjeno pravico slišal manj kot eden od štirih prebivalcev. Gre za območja, na katerih je večji delež tistih, ki jim grozi revščina. To kaže na splošno težavo, da se ljudje, ki so v družbi bolj prikrajšani, te pravice praviloma manj zavedajo. Podatki kažejo, da se te pravice redkeje zavedajo ljudje, ki ne delajo, ki imajo pogosteje težave s plačevanjem svojih računov, ki živijo na podeželju ali so starejši.

SLIKA 4: OZAVEŠČENOST O PRAVICI IZ GDPR DO ZAVRNITVE NEPOSREDNEGA TRŽENJA V EU IN ZDRUŽENEM KRALJESTVU PO DRŽAVAH IN REGIJAH (V %)



Opomba: Na zemljevidu niso prikazane države, ki niso članice EU, razen Združenega kraljestva. Svetlejše senčenje = več ljudi se zaveda svoje pravice. Temnejše senčenje = manj ljudi se zaveda svoje pravice. Rezultati v državah ali na območjih, ki so pobarvani s svetlo sivo barvo, so bili izključeni, ker je bilo anketirancev manj kot 20, kar pomeni, da je bilo število ugotovitev prenizko za pridobitev zanesljivih rezultatov. N = 26 503. Vprašanje: „Splošna uredba o varstvu podatkov (GDPR) zagotavlja številne pravice. Ste že slišali za vsako od naslednjih pravic? […] 18.2 Pravica zavrnitve neposrednega trženja.“

Vir: FRA, 2020 [Izračuni in predstavitev na podlagi podatkov Evropske komisije (2019), **Eurobarometer, 91.2***]*

Izkušnje na podlagi primerov uporabe

Na splošno so anketirani strokovnjaki poudarili, da je interpretacija prava o varstvu podatkov otežena in premalo jasna, zlasti ko govorimo o pomenu avtomatiziranega odločanja. Eden od francoskih strokovnjakov je menil, da bi zaradi težav pri pojasnjevanju morali biti vsi postopki avtomatiziranega odločanja prepovedani, kar pomeni, da bi bilo treba izjeme v GDPR, ki omogočajo avtomatizirano odločanje, odpraviti. Poudaril je, da je umetno inteligenco mogoče uporabiti zgolj kot orodje za podporo pri odločanju.

Drug strokovnjak, neodvisni odvetnik iz Nizozemske, meni, da sedanji zakoni in standardi zadostujejo, vendar pravi, da jih je treba konkretizirati po posameznih sektorjih. Zlasti poudarja, da obseg obstoječih pravil za dovoljene primere avtomatiziranega odločanja ni jasen in da ostajata nejasna tudi pojma celovita ocena in „človek v zanki“. Enaki argumenti so bili izpostavljeni tudi v povezavi s primerom SyRI, kjer ni bilo jasno, v kolikšnem obsegu so bile odločitve predmet nadzora.

Drug strokovnjak, zaposlen pri nadzornem organu, na splošno ne vidi potrebe po spremembah zakonodaje o varstvu podatkov, saj je „zakonodaja precej izčrpna. Gre bolj za organizacijo nadzora in tudi za politično voljo, ki stoji za njim“.

Ti pomisleki odražajo ugotovitve drugih raziskav, ki izpostavljajo resne pomisleke v zvezi s pravico do človeškega nadzora. Odgovorni uradniki so na primer o rezultatih algoritemskega sistema, zgrajenega za profiliranje brezposelnih na Poljskem, podvomili le v manj kot enem odstotku primerov. Na ta način se namreč podporno orodje pravzaprav prelevi v orodje avtomatiziranega odločanja[27].

Z vprašanjem pregleda odločitev ali rezultatov iz sistemov umetne inteligence je povezan tudi izziv očitnega pomanjkanja znanja o načinu delovanja umetne inteligence. Anketiranci pogosto niso mogli podrobno pojasniti, kako sistem, ki ga uporabljajo, deluje ali katere podatke uporablja, pa naj bo to zaradi pomanjkanja znanja ali premajhne preglednosti. Pomembne informacije o vključeni logiki in razlaga rezultatov algoritmov so bistvenega pomena za več temeljnih pravic. Ključnega pomena niso le za obdelavo osebnih podatkov, ampak tudi za zagotavljanje, da so algoritmi pravični in ne diskriminirajo. Ljudem je prav tako treba omogočiti, da izpodbijajo odločitve in ugovarjajo sistemom umetne inteligence.

Anketiranec, zaposlen v javni upravi, pojasnjuje, da se kompleksnost razlikuje glede na konkretne naloge. Sistemi za upravljanje licenc so lahko razmeroma enostavni. Analiza za preprečevanje kriminala uporablja več virov podatkov, kar otežuje njeno razumevanje. Drug sogovornik, zaposlen v organih pregona, pravi, da umetna inteligenca, ki jo trenutno uporabljajo policijske organizacije, še ni tako zapletena, da bi jo bilo nemogoče razumeti, vendar ni nujno, da bo to veljalo tudi v prihodnosti.

Anketiranec, ki se ukvarja s podatki o finančnih transakcijah, navaja, da so bili tradicionalni modeli preprosto razumljivi. Vendar pa je nove metodologije težje pojasniti in podjetje mora vlagati sredstva v to, da bodo modeli bolj razumljivi. Vendar pa stopnja razumljivosti, ki jo zahteva GDPR, po mnenju anketiranca ni jasno določena.

„Obstaja tveganje prevelikega zaupanja v stroje.“
(javna uprava, Francija)

„Na splošno obstaja velika napetost glede splošne uredbe o varstvu podatkov. Zelo si želimo, da bi nam šlo dobro, v resnici pa stvari morda le še poslabšamo, saj se interpretacija podatkov potem izkaže za nemogočo nalogo.“
(javna uprava, Nizozemska)

„Če bi morali model razložiti, nam ne bi uspelo. Gre za statistični model, ki ga ni mogoče dobro razložiti.“
(javna uprava, Francija)

„Interno lahko odločitve modelov strojnega učenja pojasnimo in za to imamo več sredstev.“
(zasebni sektor, Estonija)

„Če sistemi nimajo črnih skrinjic z informacijami ali procesi, je to že korak naprej v smeri varstva človekovih pravic.“
(javna uprava, Španija)

„Zelo naklonjeni smo ideji, da mora biti umetna inteligenca razložljiva.“
(javna uprava, Francija)

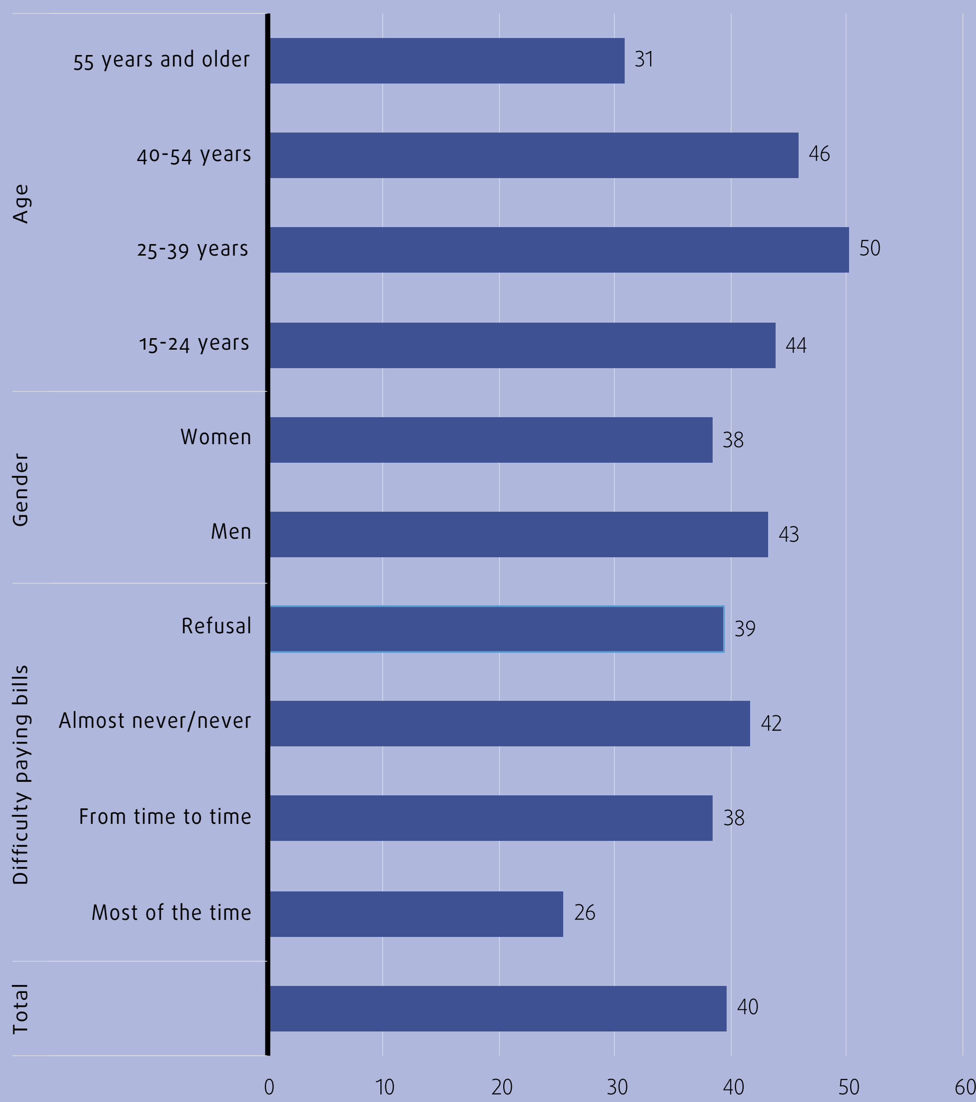
Ozaveščenost o pravici vpliva na avtomatizirano sprejemanje odločitev

Dokazi kažejo na to, da se večina ljudi ne zaveda svoje pravice vpliva na avtomatizirano sprejemanje odločitev. Raziskava Eurobarometra je pokazala, da se 40 % Evropejcev zaveda svoje pravice do varstva podatkov.

Analiza raziskave Eurobarometra, ki jo je izvedla FRA, kaže na to, da je ta odstotek pri ljudeh z nižjim socialno-ekonomskim statusom znatno nižji. Samo 26 % državljanov EU, ki so povedali, da s težavo plačujejo račune, ve za obstoj te pravice. Pomanjkljiva ozaveščenost socialno prikrajšanih bi lahko prispevala k njihovi nadaljnji socialni izključenosti, ko se že tako prikrajšani posamezniki redkeje zavedajo, da lahko izpodbijajo (avtomatizirane) odločitve o sebi (glej sliko 5).

Razlike med spoloma so majhne, vendar se ženske te pravice zavedajo še nekoliko redkeje (38 % žensk in 43 % moških). Starejši ljudje so precej manj ozaveščeni (31 % ljudi, starih 55 let ali več).

SLIKA 5: OZAVEŠČENOST O PRAVICI VPLIVA NA AVTOMATIZIRANO SPREJEMANJE ODLOČITEV PO STAROSTI, SPOLU IN GLEDE NA TEŽAVE LJUDI S PLAČEVANJEM POLOŽNIC (V %)

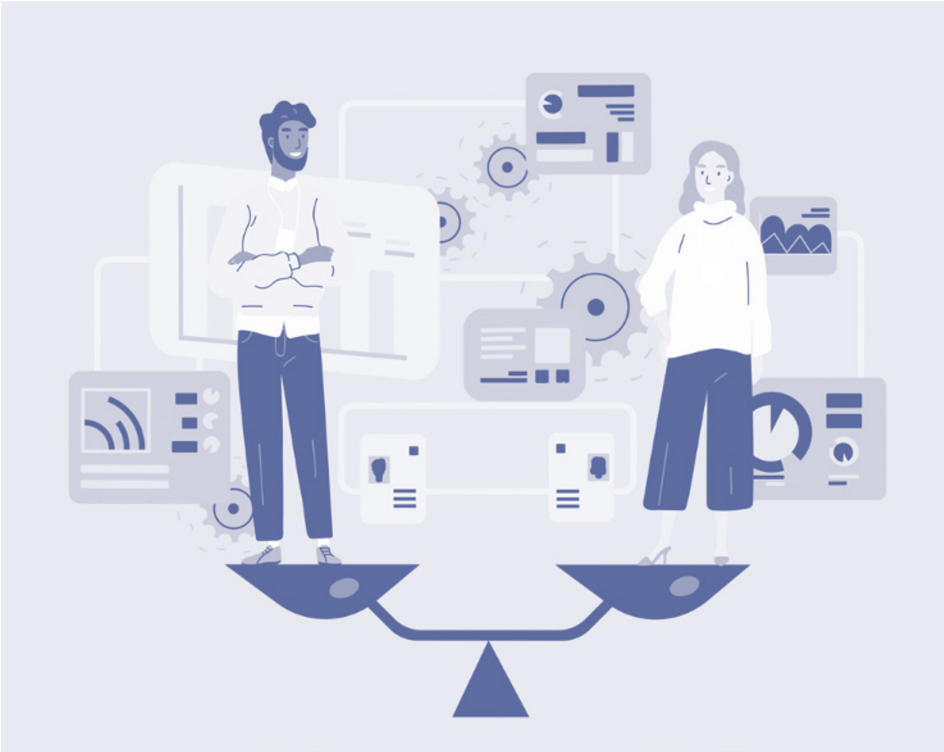


Opombe: N = 26 503. Vprašanje: „Splošna uredba o varstvu podatkov (GDPR) zagotavlja številne pravice. Ste že slišali za vsako od naslednjih pravic? […] 18.5 Pravica vpliva na avtomatizirano sprejemanje odločitev (npr. ko algoritem odloči, ali vam bo odobreno posojilo ali ne).“

Vir: FRA, 2020 [izračuni in predstavitev na podlagi podatkov Evropske komisije (2019), **Eurobarometer, 91.2***]*

4.5 ENAKOST IN PREPOVED DISKRIMINACIJE

Enakost pred zakonom in prepoved diskriminacije sta določeni v členih 20 in 21 Listine. Do diskriminacije pride, „kadar je, je bila ali bi bila oseba obravnavana manj ugodno kakor neka druga v primerljivi situaciji" na podlagi zaznane ali resnične osebne okoliščine[28] (ki jo imenujemo „zaščitena osebna okoliščina/značilnost"). Člen 21 Listine prepoveduje vsakršno diskriminacijo na podlagi spola, rase, barve kože, etničnega ali socialnega porekla, genetskih značilnosti, jezika, vere ali prepričanja, političnega ali drugega mnenja, pripadnosti narodnostni manjšini, premoženja, rojstva, invalidnosti, starosti ali spolne usmerjenosti.



Prepoved iz Listine odraža ustrezni pravici iz EKČP (člen 14) in Protokola št. 12 k EKČP (člen 12), vendar je ureditev še nekoliko širša, saj vzpostavlja neizčrpen, odprt seznam, ki varstvo pred diskriminacijo širi na spekter novih osebnih okoliščin. Za razliko od člena 14 EKČP je pravica do nediskriminacije iz Listine samostojna pravica, ki se uporablja za primere, za katere ni treba, da so zajeti v kateri drugi določbi Listine[29].

Glavni izzivi

Diskriminacija je ključna tema z vidika uporabe umetne inteligence, saj je namen algoritmov strojnega učenja ravno kategorizacija, razvrščanje in ločevanje. Kot poudarja eden od intervjuvanih strokovnjakov, razlikovanje samo po sebi ni nič slabega. Po njegovem mnenju se pri odločanju o odobritvi posojila za razlikovanje med posamezniki lahko uporabi kreditna zgodovina, ne pa zaščitene osebne okoliščine, kot sta spol ali vera. Vendar pa so številne osebne lastnosti ali življenjske izkušnje pogosto tesno povezane z zaščitenimi osebnimi okoliščinami. Kreditna zgodovina je pri moških in ženskah lahko sistematično drugačna, že zaradi razlik v zaslužkih in zgodovini delovnih mest.

Anketiranci kot glavni namen uporabe tehnologij umetne inteligence pogosto omenjajo učinkovitost. Vendar je treba opozoriti, da ta ne more upravičiti nepoštenega in neenakega obravnavanja.

Zaščitene osebne okoliščine so pogosto zelo tesno povezane s tveganji. Razlike v življenjskih razmerah med moškimi in ženskami so lahko pogosto povezane z različnimi zavarovalnimi tveganji. Vendar, kot poudarja sodba Sodišča v zadevi Test-Achats[30], to ni sprejemljivo. V tem primeru je Sodišče EU odpravilo diskriminacijo na podlagi spola pri določanju cen zavarovanja[31].

V določenih okoliščinah in na nekaterih področjih pa bi lahko uporaba algoritmov celo pozitivno prispevala k zmanjšanju pristranskosti in stereotipov. Algoritemska analiza podatkov lahko daje rezultate, ki bi lahko prispevali k razbitju predsodkov. Na primer, napovedno policijsko delo bi lahko v nekaterih okoliščinah privedlo do bolj pravičnega in manj diskriminatornega policijskega dela prek zmanjšanja odvisnosti od subjektivne človeške presoje[32]. Tehnike predvidevanja se lahko uporabijo pri preiskovanju tako imenovane kriminalitete belih ovratnikov, kot so kazniva dejanja na področju financ, ki so bila v preteklosti premalo obravnavana[33].

Kljub temu pa se neposredna ali posredna diskriminacija[34] z uporabo algoritmov, ki vsebujejo masovne podatke, šteje za enega najbolj perečih izzivov pri uporabi tehnologij umetne inteligence[35]. Pristranskost in diskriminacija, vključno z diskriminacijo na podlagi spola, se lahko pri sprejemanju podatkovno podprtih algoritemskih odločitev pojavita iz več razlogov in na številnih ravneh sistemov umetne inteligence. Njuno odkrivanje in preprečevanje je zelo oteženo[36]. Vir potencialne diskriminacije in nepoštene obravnave je pogosto povezan s kakovostjo in pristranskostjo podatkov[37].

Diskriminatorne učinke zoper nekatere skupine je v praksi zelo težko izpodbijati[38]. Do tega trenutka je bila diskriminacija v povezavi s sistemi umetne inteligence obravnavana le v omejenem številu zadev pred sodišči[39].

Pritožbeno sodišče Združenega kraljestva: policijska uporaba tehnologij za prepoznavanje obraza krši človekove pravice

S prvostopenjsko odločbo oddelčnega sodišča v Cardiffu je bil leta 2019 zavrnjen zahtevek glede zakonitosti uporabe sistema za prepoznavanje obraza AFR Locate s strani policije v Južnem Walesu. Pritožbeno sodišče je to odločitev razveljavilo.

Ugotovilo je, da je bila uporaba programa za prepoznavanje obraza s strani policije nezakonita. Pritožbeno sodišče je razsodilo, da je „trenutno posameznim policistom prepuščena prevelika diskrecija“. Dodalo je, da „ni jasno, kdo je lahko na seznamu opazovanih oseb in prav tako ni jasno, kakšna so merila za določanje, kje se [tehnologija] lahko uporabi“*.

Sodišče je hkrati ugotovilo, da policija ni opravila zadostnih raziskav za ugotavljanje, ali se je uporabljena programska oprema izkazala za pristransko na podlagi rase ali spola.

Ta sodba je prva v Evropi, ki posebej obravnava ta vprašanja. Znatno zmanjšuje obseg tega, kaj je dovoljeno in kaj morajo storiti organi kazenskega pregona, da bi v celoti zagotovili spoštovanje zakonodaje o človekovih pravicah**.

** Združeno kraljestvo, Pritožbeno sodišče,* ***R (Bridges) proti CC South Wales****, [2020] EWCA Civ 1058, 11. avgust 2020.*

*** Ars Technica,* **Police use of facial recognition violates human rights, UK court rules** *(Sodišče v Združenem kraljestvu je odločilo, da policijska uporaba tehnologije za prepoznavanje obraza krši človekove pravice), 11. avgust 2020.*

Študije so poudarile možnost diskriminacije zaradi uporabe sistemov umetne inteligence na področjih, ki jih obravnava poročilo[40]. Na področju napovednega policijskega dela obstaja denimo posebno tveganje, povezano z morebitnimi orodji za avtomatizirano odločanje, za reproduciranje in utrjevanje obstoječih diskriminatornih praks, ki spodkopavajo enakost pred zakonom (člen 20 Listine). Zgodovinski podatki o kaznivih dejanjih, na katerih temelji napovedno policijsko delo, so lahko pristranski[41] in odražajo določene vrzeli v podatkih (npr. kronično neprijavljanje nekaterih vrst kaznivih dejanj) ali težave pri zapisovanju podatkov (npr. človeške napake, vendar tudi pristranskost posameznih uradnikov).

Raziskave o žrtvah kaznivih dejanj dosledno kažejo, da javnost velikega deleža kaznivih dejanj nikoli ne prijavi policiji, kar velja zlasti za kazniva dejanja, ki vključujejo fizično in/ali spolno nasilje, ter kazniva dejanja iz sovraštva. Raziskava FRA o nasilju nad ženskami, v kateri je sodelovalo 42 000 anketirank, je na primer pokazala, da je samo ena od petih žensk, ki so doživele nasilje s strani partnerja ali kogar koli drugega, najhujši incident prijavila policiji[42]. Raziskava FRA EU-MIDIS II, v kateri je sodelovalo 25 500 anketirancev iz celotne Evropske unije, je pokazala, da so bili policiji ali kateri drugi organizaciji prijavljeni le trije od desetih primerov rasno motiviranih kaznivih dejanj iz sovraštva[43].

Zlasti v razvitih državah je pri kaznivih dejanjih zoper premoženje, kot so vlomi, stopnja prijav policiji višja kot pri kaznivih dejanjih iz sovraštva. To je lahko zato, ker je prijava pogoj pri zahtevkih iz naslova zavarovanja.

Če povzamemo, zanašanje na uradne statistične podatke o kaznivih dejanjih (ki temeljijo na prijavljenih kaznivih dejanjih) pri razvoju modelov umetne inteligence na področju napovednega policijskega dela je še posebej problematično pri nekaterih konkretnih kaznivih dejanjih in konkretnih skupinah.

Nekatere spremenljivke, ki se uporabljajo pri modeliranju s pomočjo umetne inteligence, so lahko povezane z raso, etnično pripadnostjo, spolom ali drugimi zaščitenimi osebnimi okoliščinami. Kompleksnost algoritmov otežuje prepoznavanje in odstranjevanje takšnih oblik pristranskosti. Namesto objektivne analize se lahko napovedno policijsko delo spremeni v mehanizem za utrjevanje obstoječih sistemskih pomanjkljivosti in krivic, ki bo ožigosan z navidezno znanstveno legitimnostjo[44].

Zaradi uporabe napovednega policijskega dela so lahko odzivi organov kazenskega pregona manj pravični, ker nekaterim kaznivim dejanjem ali področjem posvečajo več pozornosti[45]. Napovedno policijsko delo se trenutno osredotoča na kazniva dejanja zoper premoženje, kot so kraje in vlomi, ki so pogosto povezana z določeno demografijo in konkretnimi soseskami. Posledica tega je lahko, da so nekatere demografske skupine in soseske – in posamezniki, ki v njih živijo – dodatno stigmatizirani[46]. Po drugi strani je kriminal belih ovratnikov, ki ga praviloma zagrešijo posamezniki iz drugih demografskih skupin, obravnavan manj prednostno[47]. Ti vzorci policijskega nadzora, pri katerih so nekatere soseske ali skupnosti pod nesorazmerno večjim nadzorom, segajo dlje v preteklost kot uporaba umetne inteligence. Vendar je treba v praksi preveriti „zagotovilo“, da je umetna inteligenca bolj „objektivna“ in se lahko uporabi za preprečevanje diskriminacije pri policijskem delu.

Raziskovalka z Univerze v Oxfordu Sandra Wachter poudarja, da tudi v ciljno usmerjenem oglaševanju lahko pride do diskriminacije zaradi informacij, povezanih z zaščitenimi osebnimi okoliščinami. Na novo ustvarjeni profili za namene oglaševanja bi lahko pomenili posredno diskriminacijo in lahko celo zahtevali nove vidike protidiskriminacijske zakonodaje ter širitev njenega področja uporabe[48].

Izkušnje na podlagi primerov uporabe

Številni anketiranci so povedali, da lahko uporaba umetne inteligence na splošno vodi v diskriminacijo, sistemi, ki jih oni konkretno uporabljajo, pa ne. Mnogi so izrazili prepričanje, da je izključitev informacij o zaščitenih osebnih okoliščinah lahko zadostna zaščita pred diskriminacijo. Vendar lahko pride do diskriminacije zaradi drugih informacij v nizih podatkov, ki lahko kažejo na zaščitene osebne okoliščine. Sledovi zaščitenih osebnih okoliščin so pogosto skriti tudi v drugih informacijah.

Primer javnega organa, ki uporablja umetno inteligenco pri obdavčitvah in carini, kaže na izzive, povezane z ugotavljanjem morebitnih pristranskosti in diskriminacije pri uporabi algoritmov. Med natančnim preučevanjem uporabljenih algoritmov je organ javne uprave odkril večjo stopnjo napak v davčnih napovedih posameznikov, katerih davčne številke so bile izdane šele pred kratkim, pri čemer so bile te v veliki večini primerov izdane migrantom. To je spodbudilo nadaljnje raziskave o korelaciji. Izkazalo se je, da so vloge oseb, katerih davčne številke so bile izdane šele pred kratkim, pogosteje vsebovale napake, saj ti posamezniki nikoli prej niso vlagali davčnih napovedi in niso vedeli, kako naj to storijo (kar je veljalo tudi za tiste, ki niso nujno bili priseljenci). To je tudi primer povezanih informacij, kjer bi deli številke lahko kazali na status priseljenca.

Drug anketiranec, ki je preučeval morebitno uporabo umetne inteligence za odkrivanje goljufij v zvezi s socialnimi prejemki, je v zvezi s tem povedal: „Če želite na primer preprečiti diskriminacijo na podlagi etnične pripadnosti, ne zadostuje, da preprosto odstranite oznako etnične pripadnosti, saj je tudi soseska posreden pokazatelj etnične pripadnosti oziroma igra etnična pripadnost pri tem določeno vlogo. Tako [preprečevanje diskriminacije] pogosto presega ‚neposredne' osebne okoliščine."

„Če želite, da stroj ne diskriminira na podlagi spola, preprosto ne vključujte spremenljivke spola ali naredite primere simetrične, če opazite, da ima spol določen pomen."
(javna uprava, Španija)

Četudi se je večina anketirancev zavedala splošnega potenciala za diskriminacijo pri uporabi umetne inteligence, so možnost, da je njihov sistem diskriminatoren na podlagi zaščitenih okoliščin, pogosto povsem izključili. Nekateri anketiranci so bili mnenja, da imajo njihova orodja pozitiven učinek v smislu varstva pred diskriminacijo. Eden od anketirancev, ki se ukvarja s preizkušanjem umetne inteligence na področju odločitev o socialnih prejemkih, obžaluje, da zaradi varstva podatkov umetne inteligence ne more uporabljati, saj bi po njegovem mnenju avtomatizacija omogočila učinkovito obdelavo velikih podatkovnih nizov brez kakršne koli diskriminacije. Ob ugotovitvi, da je treba upoštevati varstvo osebnih podatkov, anketiranec meni, da to ovira takojšnje sprejemanje odločitev in nediskriminacijo – „če je nek postopek mogoče avtomatizirati, bi ga bilo treba avtomatizirati".

„Če nimate dostopa do občutljivih osebnih podatkov, je nemogoče preveriti, ali na podlagi teh podatkov poteka profiliranje."
(javna uprava, Nizozemska)

Nekateri anketiranci niso jasno povedali oziroma niso bili prepričani, ali je njihova uporaba umetne inteligence lahko diskriminatorna. Prav tako so anketiranci večkrat navedli, da njihov sistem ne more diskriminirati, ker ne vključuje podatkov o zaščitenih osebnih okoliščinah. Več anketirancev s področja napovednega policijskega dela in kazenskega pregona je denimo povedalo, da možnosti za diskriminacijo ni, saj sistemi umetne inteligence ne uporabljajo niti ne ustvarjajo podatkov o zaščitenih osebnih okoliščinah in prav tako niso namenjeni identifikaciji ljudi.

Drugi, ki so se prav tako ukvarjali z napovednim policijskim delom, so menili, da bi se diskriminacija lahko pojavila, in to zlasti zaradi težav v učnih podatkih. V zvezi s primerom, pri katerem so bili v okviru napovednega policijskega dela oblikovani toplotni zemljevidi, je eden od anketirancev navedel, da ker nabor podatkov nikoli ni v celoti nevtralen, reprezentativen ali popoln, obstaja veliko tveganje pristranskosti in morebitne diskriminacije določenih skupin. Kot enega od načinov za zmanjšanje tega tveganja je predlagal izmenjavo naborov podatkov z namenom povečanja količine razpoložljivih podatkov, vendar to po njegovem mnenju v praksi onemogočajo predpisi o varstvu podatkov. Predlagal je tudi ustanovitev skupin na več ravneh, ki bi bile zadolžene, da obiskujejo različne policijske organe in preverjajo kakovost uporabljenih sistemov.

Na področju ciljno usmerjenega oglaševanja so diskriminacijo kot morebitno težavo anketiranci omenjali zlasti potem, ko so bili neposredno vprašani o tej temi. Na splošno anketiranci menijo, da njihovi sistemi ne diskriminirajo. Trije anketiranci so povedali, da se informacije o spolu in starosti ne uporabljajo in posledično v zvezi s tem ne more priti do diskriminacije. Drug anketiranec ni vedel, ali so ti podatki vključeni ali ne.

„Pri diskriminaciji se nekoliko zaplete, ker se nekatere bolezni pogosteje pojavljajo pri določenih etničnih skupinah. Napovedi tako upoštevajo spol, etnično pripadnost in genske značilnosti. Vendar to še ne pomeni diskriminacije ali kršitve človekovih pravic.“
(zasebni sektor, Francija)

Anketiranec, ki je delal na orodju za odkrivanje raka dojke, je poudaril, da so starost, spol in etnična pripadnost pomembni dejavniki, saj se pri nekaterih skupinah prebivalstva določene oblike raka pojavljajo pogosteje. Anketiranci, zaposleni v zdravstvu, so poudarili, da so možnosti za diskriminacijo povezane tudi s tem, kdo sistem uporablja, in navedli, da bi to lahko postal večji izziv, če bi sistem uporabljalo nemedicinsko osebje.

Drugačen, a povezan primer prihaja od anketiranca, ki dela na področju ocenjevanja kreditne sposobnosti v zasebnem podjetju, ki s pomočjo algoritmov odplačno pripravlja ocene kreditne sposobnosti posameznikov. Družba v svojih modelih ocenjevanja kreditne sposobnosti uporablja informacije o spolu, starosti in državljanstvu. Te informacije imajo določen vpliv na oceno kreditne sposobnosti. Na primer, mlajši ljudje ali nedržavljani imajo višjo oceno kreditnega tveganja, vendar je vpliv demografije v primerjavi s podatki iz kreditne zgodovine veliko manjši. Po mnenju anketiranca njihov sistem „zagotovo ne vpliva na pravico do nediskriminacije, saj ne sprejemajo nobenih odločitev, temveč samo prodajajo podatke in analizo podatkov. Upniki so tisti, ki morajo paziti, da ne prihaja do diskriminacije“.

Drug anketiranec, ki se ukvarja s podatkovno strategijo za finančno institucijo v zasebnem sektorju in uporablja umetno inteligenco za analizo finančnih transakcij, jasno omenja izzive, povezane z razumevanjem tega, kaj pomeni diskriminacija pri tem delu. Izpostavlja na primer, da ni jasno, v kolikšni meri je nezakonita izključitev starejših v primeru odobritve posojila, če je njihova pričakovana življenjska doba krajša od dobe odplačevanja hipoteke, ki so jo zahtevali.

Te ugotovitve kažejo na negotovost in dvom v finančnem sektorju glede tega, kako se člen 21 Listine o prepovedi diskriminacije uporablja v resničnih življenjskih situacijah[49].

Ranljive skupine

Večina razprav in raziskav o diskriminaciji pri uporabi umetne inteligence je povezanih s pristranskimi rezultati glede na etnično poreklo, spol in do neke mere tudi starost. Čeprav je pomembno analizirati morebitno diskriminacijo teh skupin, Listina zajema veliko več osebnih okoliščin, ki zahtevajo zaščito pred diskriminacijo in ki jim je posvečen le manjši del razprav in raziskav.

Te druge osebne okoliščine vključujejo na primer politično prepričanje, spolno usmerjenost in invalidnost. Listina nekaterim posebnim skupinam (poleg tistih iz členov 20 in 21) zagotavlja posebne pravice, vključno s pravicami otroka (člen 24), pravicami starejših (člen 25) in pravicami invalidov (člen 26).

Vprašanje starosti (v zvezi s starejšimi starostnimi skupinami in mlajšimi odraslimi) je med razgovori prišlo na plan zlasti v zvezi z zavarovanjem in posojili (glej zgoraj).

Vendar nihče od anketirancev ali strokovnjakov ni neposredno omenil pravic otroka. To je lahko v določeni meri povezano z naravo primerov uporabe, ki so bili predmet raziskave, vendar ta podatek odraža tudi dejstvo, da omenjena tema ne sodi med prednostne teme posameznikov, ki se ukvarjajo z umetno inteligenco.

Člen 24 Listine poudarja, da je pri vseh ukrepih javnih organov ali zasebnih ustanov, ki se nanašajo na otroke, treba upoštevati predvsem koristi otroka, kar seveda na enak način velja tudi za področje umetne inteligence[50].

Možno uporabo umetne inteligence na področju skrbništva otrok in razporejanja otrok po šolah sta omenila le dva anketiranca, zaposlena v javni upravi. Vendar te teme nista omenila v povezavi z upoštevanjem otrokovih koristi. Pravzaprav anketiranca nista želela podrobneje razpravljati o teh primerih uporabe, kar bi lahko odražalo občutljivost te teme.

Nazadnje, v nobenem od razgovorov niso bila omenjena vprašanja, povezana z vključevanjem invalidov.

Ozaveščenost splošne populacije o možnostih, da bi umetna inteligenca privedla do diskriminacije

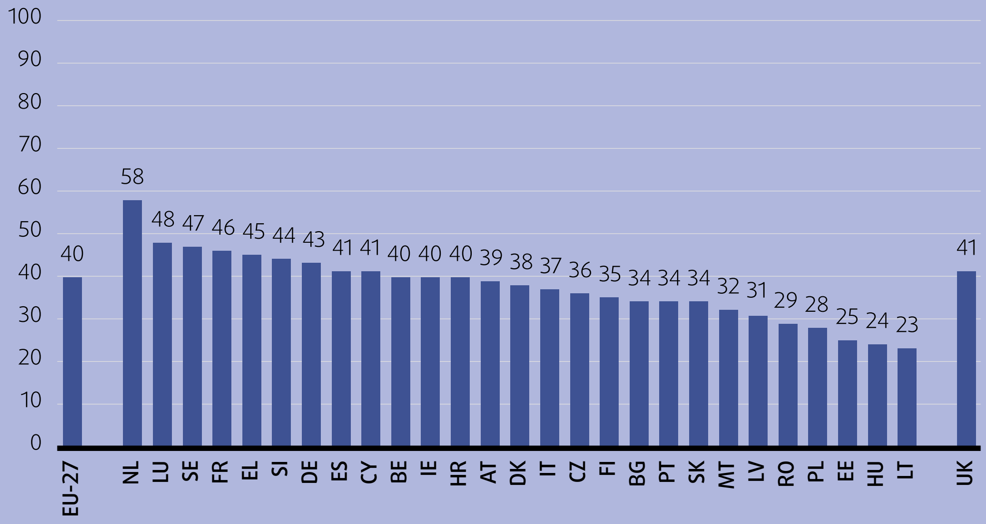
V okviru raziskave Eurobarometra, ki je vključevala vprašanja o umetni inteligenci, so bili anketiranci vprašani o področjih, ki jih z vidika umetne inteligence najbolj skrbijo, vključno z diskriminacijo pri odločanju, nejasno odgovornostjo in tem, da se ni mogoče nikomur pritožiti.

Le približno 40 % državljanov EU je navedlo, da jih skrbi, da bi uporaba umetne inteligence lahko vodila do diskriminacije na podlagi starosti, spola, rase ali državljanstva – na primer pri sprejemanju odločitev o zaposlovanju, kreditni sposobnosti itd.

Rezultati se v posameznih državah razlikujejo. Diskriminacija skrbi večji delež prebivalstva na Nizozemskem (58 %), v Luksemburgu (48 %) in na Švedskem (47 %). Nižji odstotek prebivalstva je zaradi tovrstnih tem zaskrbljen v Estoniji (25 %), na Madžarskem (24 %) in v Litvi (23 %) (glej sliko 6).

Vendar iz zastavljenega vprašanja ni jasno, ali ljudje ne vedo, da lahko pride do diskriminacije, ali se te možnosti zavedajo in hkrati menijo, da ni problematična.

SLIKA 6: OZAVEŠČENOST O TVEGANJIH ZA POJAV DISKRIMINACIJE PRI UPORABI UMETNE INTELIGENCE V POSAMEZNIH DRŽAVAH (V %)



Opombe: Vključeni so posamezniki, ki so odgovorili, da jih skrbi možnost, da bi umetna inteligenca vodila do diskriminacije pri enem od treh vprašanj ali pri vseh treh vprašanjih.

Vir: Izračuni FRA na podlagi podatkov Evropske komisije (2019), ***Eurobarometer, 92.3***

Obravnava vprašanja neenakosti spolov pri zasnovi in uporabi umetne inteligence

Listina določa, da se mora enakost žensk in moških zagotoviti na vseh področjih, vključno z zaposlovanjem, delom in plačilom za delo (člen 23). Diskriminacija na podlagi spola je eno od glavnih vprašanj pri zasnovi in uporabi umetne inteligence in sorodnih tehnologij*.

Z vidika razvoja Evropski ekonomsko-socialni odbor ugotavlja, da se umetna inteligenca razvija v homogenem okolju, ki ga povečini sestavljajo mladi belci. Posledica tega so kulturne razlike in razlike med spoloma, ki se vnašajo v tehnologije umetne inteligence. Učni podatki so na primer izpostavljeni prikrajanju, so lahko tendenciozni, odražati utegnejo kulturne, spolne ali druge preference ali predsodke, lahko pa so celo napačni**. To se odraža tudi v tej raziskavi, v kateri je bila kljub prizadevanjem za doseganje uravnoteženosti spolov večina anketirancev moških.

Razlike v fazi oblikovanja in uvajanja so povezane s sistematično prikrajšanostjo, ki so ji izpostavljene ženske na trgu dela, in morebitnim pomanjkanjem ozaveščenosti o pristranskosti glede na spol. Nedavna študija je pokazala, da bi lahko povečana uporaba industrijskih robotov povečala tudi razlike med spoloma, kljub temu da bi oba spola imela koristi od večje avtomatizacije, saj je analiza pokazala, da bi imeli moški v srednje- in visokokvalificiranih poklicih nesorazmerne prednosti***.

Kar zadeva prihodnost, bi lahko uporaba podatkov in algoritmov pripomogla k boljšemu vključevanju enakosti spolov v politike in procese prek posvečanja pozornosti spolno zaznamovanim naborom podatkov. Razprave o neenakostih med spoloma in uporabi podatkov („podatkovni feminizem")**** bi lahko pripomogle k ozaveščanju, da se stališča moških ne bi smela obravnavati kot privzeta, kar nato vpliva tudi na zbirke podatkov*****.

* *Glej tudi Evropska komisija,* Bela knjiga o umetni inteligenci – evropski pristop k odličnosti in zaupanju, *COM(2020) 65 final, Bruselj, 19. februar 2020, str. 1.*

** *Evropski ekonomsko-socialni odbor, Umetna inteligenca – posledice za enotni (digitalni) trg, proizvodnjo, potrošnjo, zaposlovanje in družbo (mnenje na lastno pobudo), 31. maj 2017, UL C 288, str. 1.*

*** *Aksoy, C., Özcan, B. in Philipp, J. (2020),* **Robots and the Gender Pay Gap in Europe** *(Roboti in razlike v plačilu med spoloma v Evropi), dokument za razpravo IZA št. 13482.*

**** *Glej* ***spletno stran o podatkovnem feminizmu*** *na spletnem mestu organizacije Datasociety.*

***** *Criado Perez, C. (2020),* Invisible Women. Exposing data bias in a world designed for men *(Nevidne ženske. Razkrivanje podatkovne pristranskosti v svetu, ki je zasnovan za moške), London.*

4.6 DOSTOP DO PRAVNEGA VARSTVA

Pravica do učinkovitega pravnega sredstva in nepristranskega sodišča (člen 47 Listine) je ena od najpogosteje uporabljenih pravic iz Listine v sodnih postopkih. To kaže na pomen ohranjanja temeljnih pravic in pravne države. Ta pravica horizontalnega značaja posameznikom omogoča, da izpodbijajo ukrep, ki vpliva na katero koli njihovo pravico, ki jim jo podeljuje pravo EU, ne samo na pravice, ki jih zagotavlja Listina[51]. Sodišče EU je poudarilo, da člen 47 Listine ponovno potrjuje načelo učinkovitega sodnega varstva in da je treba značilnosti pravnega sredstva določiti na način, ki je skladen s tem načelom[52].

Pravica do učinkovitega pravnega sredstva zajema tudi odločitve, sprejete s podporo tehnologij umetne inteligence. Zakonodaja EU na področju varstva podatkov ponovno poudarja, da je treba zagotoviti pravico do učinkovitega pravnega sredstva v zvezi z odločitvami upravljavca ali obdelovalca[53] ter nadzornega organa[54]. Podatki, ki jih obdelujejo tehnologije umetne inteligence, niso izjema.

Pomembno je opozoriti, da se možnost vložitve upravne pritožbe pri nadzornem organu, kot jo določata GDPR in direktiva o kazenskem pregonu[55], ne šteje za učinkovito pravno sredstvo v skladu s členom 47 Listine. To je zato, ker v presojo v tem primeru ni vključeno nobeno sodišče. Sodno varstvo mora vedno biti razpoložljivo in dostopno, če se notranji mehanizmi in mehanizmi alternativnega reševanja sporov izkažejo za nezadostne ali če se zadevna oseba odloči za sodno varstvo[56].

Uporaba umetne inteligence lahko predstavlja izziv z vidika pravice do učinkovitega pravnega sredstva, in sicer na več načinov. Pomembna skrb je pomanjkanje preglednosti pri uporabi in delovanju novih tehnologij. Algoritemsko odločanje je notorično nepregledno: zbiranje podatkov, učenje algoritmov, izbira podatkov za modeliranje ali profiliranje, situacija v zvezi s posameznikovo privolitvijo, učinkovitost in stopnje napak algoritma ter drugi vidiki pogosto niso pregledno predstavljeni[57].

Brez dostopa do teh informacij se posamezniki morda ne bodo mogli braniti, odločiti, kdo je odgovoren za odločitve, ki jih zadevajo[58], se pritožiti na katero koli odločitev, ki negativno vpliva nanje, ali imeti zagotovljeno pravico do poštenega sojenja, ki vključuje načelo enakosti orožij ali kontradiktornega postopka in jo določa tudi ESČP[59]. Te zahteve so v skladu s členom 52(3) Listine prav tako del ustrezne pravice iz Listine (člen 47).

Glavni izzivi

Ta vprašanja se odražajo v posebnih izzivih na področju pravice do učinkovitega pravnega sredstva in poštenega sojenja, ki so jih izpostavili strokovnjaki, s katerimi so se pogovarjali. Na splošno strokovnjaki poudarjajo razliko pri dostopu do pravnih sredstev v zasebnih podjetjih in javni upravi. Javni organi so pogosteje prisiljeni, da pri uporabi umetne inteligence ravnajo pregledno. Medtem se zdi, da so podjetja bolj skrivnostna, kar nakazuje ocena več strokovnjakov. Vendar pa je nizozemski strokovnjak dejal, da bi lahko bili ljudje bolj pripravljeni pritožiti se zoper odločitev podjetja kot pa nasprotovati odločitvi javnega organa. To je zato, ker javne storitve pogosto zadevajo ranljive posameznike, ki potrebujejo socialne ugodnosti in so manj nagnjeni k pritožbam zoper kakršne koli odločitve.

Možnost uspešne pritožbe zoper uporabo umetne inteligence in izpodbijanja odločitve, sprejete na podlagi umetne inteligence, je bistvenega pomena z vidika dostopa do sodnega varstva. V intervjujih je bilo v zvezi s tem izpostavljeno naslednje:

- ozaveščanje ljudi, da se uporablja umetna inteligenca,
- ozaveščanje ljudi o tem, kako in kje se lahko pritožijo,
- zagotavljanje, da je sisteme umetne inteligence in odločitve, ki temeljijo na umetni inteligenci, mogoče razložiti.

Najprej morajo vsi vedeti, ali imajo opravka s sistemom umetne inteligence. Če sprejeta odločitev vpliva na ljudi, npr. na njihove socialne prejemke, se lahko ti zoper odločitev načeloma pritožijo, vendar se ne bodo mogli pritožiti zoper dejstvo, da se uporablja umetna inteligenca, če o tem niso obveščeni.

Eden od strokovnjakov je pojasnil, da je kljub splošni volji po pritožbi največja težava v tem, da ljudje pogosto sploh ne vedo, da se uporablja umetna inteligenca, saj organizacije pri tem ne ravnajo pregledno, čeprav jim to GDPR nalaga. Več anketirancev je navedlo, da je obveščanje ljudi o tem, da vsaka odločitev o njih temelji na (delno) avtomatiziranih orodjih, prvi korak za zagotavljanje dostopa do pritožbenih mehanizmov.

Drugič, vsakdo mora vedeti, kje in kako se lahko pritoži. Ljudje zelo težko ugotavljajo, kateri organ je pristojen za obravnavo katerih pritožb. Eden od strokovnjakov je poudaril, da se potrošniki pogosto ne znajo pritožiti – na primer banki, ki bi lahko uporabila algoritme za odločanje o finančnih zadevah. Organ javne uprave, ki izdaja avtomatizirane odločitve, se je odločil dopolniti odločbe z imeni zaposlenih, s katerimi lahko posameznik, ki bi morebiti želel izpodbijati (avtomatizirano) odločitev, stopi v stik. Večina anketirancev je navedla, da obstajajo načini in postopki za vlaganje pritožb, ki so enaki kot pri postopkih, ki niso povezani z uporabo umetne inteligence. Le nekaj podjetij ali organizacij, ki uporabljajo umetno inteligenco v povezavi z anonimiziranimi ali zbirnimi podatki, navaja, da nimajo vzpostavljenih nobenih pritožbenih mehanizmov.

Nazadnje, tisti, ki želi oddati pritožbo, potrebuje dovolj informacij, na podlagi katerih lahko izpodbija zadevno odločitev. Samo natančne informacije o sistemih umetne inteligence lahko zagotovijo enakost orožij, ki omogoča smiselno izpodbijanje odločitev. Vendar je to pri uporabi umetne inteligence lahko zapleteno zaradi:

- morebitnih vprašanj v zvezi s pravicami intelektualne lastnine in
- ker je zapletene sisteme težko razložiti.

Pravice intelektualne lastnine predstavljajo eno od ovir za zagotavljanje zadostne količine informacij o načinu sprejetja odločitve ali delovanja sistema. Algoritmi so lahko del izvedene programske opreme ali tehničnega izuma, za katerega lahko veljajo pravice intelektualne lastnine – pravice, zaščitene v skladu s členom 17(2) Listine. Posamezni akterji si pogosto prizadevajo za zaščito avtorskih pravic, patentov in poslovnih skrivnosti, da bi zaščitili svoje znanje o umetni inteligenci[60].

Eden izmed anketirancev iz zavarovalniškega sektorja trdi, da se zaradi zelo konkurenčnega trga „ne sme preveč govoriti o delovanju uporabljene tehnologije", na primer o tem, zakaj je bila stranki ponujena določena cena. To je predvsem zato, ker bi lahko konkurenti izkoristili znanje, ki bi izhajalo iz pregledov osnovne programske opreme. Drug anketiranec, ki uporablja umetno inteligenco za obravnavanje vlog za izdajo vizumov, ugotavlja, da bi lahko uporaba sistemov, ki so jih razvili zunanji ponudniki in katerih algoritmi so zajeti v pravicah intelektualne lastnine, v nadaljevanju ovirala potrebno preglednost.

Drug izziv, ki stoji na poti uspešnim pritožbam zoper avtomatizirane odločitve ali splošno uporabo umetne inteligence, predstavljajo **težave pri pojasnjevanju odločitev, ki temeljijo na zapletenih sistemih**. Anketiranci, zaposleni v javni upravi, poudarjajo, da so praviloma na voljo jasna navodila glede tega, kako se pritožiti zoper upravno odločbo, gre namreč za področje, ki po mnenju anketirancev zahteva podrobna pojasnila. Na primer, pri sistemih, ki samodejno določajo nadomestila za brezposelnost v primerih, ki ne vključujejo diskrecijske pravice, lahko stranke zahtevajo obrazložitev avtomatizirane upravne odločbe. Anketiranec je v zvezi s tem navedel, da lahko stranke, če želijo vpogledati v izračune v ozadju finančnih odločitev, to storijo v samopostrežnem sistemu na spletni strani organizacije ali v izdani dokumentaciji, na obeh mestih je na voljo podroben opis uporabljenih izračunov.

Anketiranci priznavajo, da je odprta in pregledna logika bistvenega pomena za pojasnjevanje odločitev, ki temeljijo na umetni inteligenci, vendar je ta cilj pogosto težko dosegljiv ali celo neuresničljiv. Eden od anketirancev, zaposlen v banki, omenja, da kompleksnejših rešitev strojnega učenja pri nekaterih vrstah odločanja ni mogoče uporabiti, ker sklepanja sistema ni mogoče enostavno pojasniti, zato se takšni sistemi uporabljajo samo za druge namene. Vendar pa anketiranec, zaposlen v drugi banki, navaja, da se taki sistemi kljub temu uporabljajo, vendar poleg zapletenih uporabljajo tudi preprostejše metode, da bi dobili predstavo o verjetnih razlogih za odločitve.

Eden od strokovnjakov je izpostavil težavo, da podjetja sama niti interno nimajo dovolj informacij o tem, kako delujejo njihovi algoritmi. Zdi se, da predstavlja pomanjkanje strokovnega znanja in izkušenj v praksi veliko oviro pri iskanju dostopa do učinkovitega pravnega sredstva[61].

Izkušnje na podlagi primerov uporabe

Anketiranci, ki so razpravljali o orodjih za **napovedno policijsko delo**, so poudarili, da je preglednost pomembna.

V primeru uporabe, kjer je bilo govora o nasilju na podlagi spola, so menili, da bi se dalo preglednost povečati denimo s predložitvijo policijskega dosjeja in rezultatov sistema umetne inteligence sodniku ter obveščanjem žrtve o stopnji tveganja, ki je povezana s primerom, in o policijskih ukrepih, ki se bodo posledično uporabljali.

„Vprašanje preglednosti je danes zelo pomembno, obstaja veliko postopkov za objavo informacij, veliko avtomatskih sredstev, ki pomagajo pri nalaganju informacij na portale, in veliko truda je bilo vloženega v aktivnosti za povečanje preglednosti."
(javna uprava, Španija)

Anketiranci, ki so govorili o primeru s toplotnim zemljevidom, so omenjali številne zahteve za pojasnila, ki jih je policija prejela glede namena sistema in njegovega delovanja, ter izpostavili preglednost kot način za zmanjšanje zaskrbljenosti javnosti.

Številni anketiranci so opozorili na možnost, da posamezniki, ki jih sistem zadeva, vložijo pritožbo na policijo, sodišča ali institucijo varuha človekovih pravic. V zvezi s primerom nasilja v družini pa je anketiranec navedel, da ne obstaja postopek, s katerim bi bilo mogoče oporekati sistemu policijskega protokola.

V zvezi z ukrepi za varstvo temeljnih pravic v primerih uporabe, ki so se nanašali na **zdravstvene storitve**, je več anketirancev omenilo etične odbore ter splošne pravne zaščitne ukrepe in pravila o varstvu podatkov. V prvi vrsti so bili omenjeni pregledi in kontrole, ki so se izvajali prek zunanjih akterjev. V organizacijah tistih anketirancev, ki so odgovarjali na to vprašanje, niso bili vzpostavljeni nobeni posebni pritožbeni postopki.

Nekateri anketiranci so poudarili, da odgovornost za odločitve na koncu prevzamejo zdravniki in da pacienti pogosto sploh ne vedo za uporabo orodja umetne inteligence. Na primer, pri sistemih za odkrivanje raka dojke je anketiranec navedel, da ni možnih pravnih sredstev zoper razvijalca orodja, saj se za diagnozo odloča radiolog, ki je prav tako odgovoren za kakršne koli napake.

Zaščitni ukrepi pri **ciljno usmerjenem oglaševanju** večinoma sledijo zahtevam glede varstva podatkov, kot je zagotavljanje pridobitve in spoštovanja privolitve posameznikov. Eno od podjetij pa denimo skrbi, da se njegove stranke ne ukvarjajo z nezakonitimi praksami, in zavrača stranke iz nekaterih sektorjev, kot je npr. politično oglaševanje.

Prejete pritožbe

Le malo sodelujočih organizacij je že prejelo pritožbe, ki bi izpodbijale njihovo uporabo umetne inteligence. V nekaterih primerih anketiranci navajajo, da so prejeli pritožbe pritožnikov, ki niso vedeli, da je bila uporabljena umetna inteligenca, in so pri odločanju opazili nepravilne rezultate.

Posamezniki so se na primer pritožili v zvezi s prometnimi globami, pri čemer je policist ustavil voznika avtomobila in po pojasnilu voznika avtomobila, da je globa neupravičena, ročno popravil informacije v sistemu, ne da bi mu pri tem uspelo posodobiti tudi pretekle podatke v sistemu. V teh primerih namreč globe ostanejo vidne v celotnem sistemu in bo ta oseba ob vsaki priložnosti še naprej profilirana kot visoko tvegana.

„Število pritožb glede uporabe podatkov je majhno, namesto tega so ljudje večkrat zahtevali, da se izbrišejo nekatere informacije o njih.“

(zasebno podjetje, Estonija)

Čeprav organizacije le redko prejemajo uradne pritožbe v zvezi z uporabo umetne inteligence, anketiranci pogosto navajajo, da je vzrok v tem, da je njihovo uvajanje umetne inteligence šele v zgodnji fazi. Kljub temu so anketiranci poročali o ponavljajočih se zahtevah za dostop do osebnih podatkov ali njihov popravek, nekateri ljudje so prav tako zahtevali, da se njihovi podatki odstranijo, ali so zaprosili za pojasnilo, zakaj je bilo dano določeno priporočilo.

Večina anketirancev trdi, da so postopki enaki, kot če bi bila odločitev obdelana ali sprejeta s strani človeka. Po drugi strani pa je nekaj anketirancev pokazalo zanimanje za odprtje novih kanalov, ki bi omogočili analizo, razlago in pravna sredstva, ki bi zadevali uporabo njihovih rešitev na področju umetne inteligence.

Na druge pravice, povezane z dostopom do pravnega varstva iz Listine, vpliva zlasti uporaba umetne inteligence pri kazenskem pregonu. Ti vključujejo na primer domnevo nedolžnosti (člen 48 Listine).

Pri ugotavljanju identitete oseb, ki so osumljene storitve kaznivega dejanja, lahko policija svoje dejavnosti usmeri konkretno proti eni osebi ali jo postavi pod sum zgolj na podlagi pomanjkljivih in razdrobljenih podatkov ter algoritemskega profiliranja[62]. Nekritično zanašanje na avtomatizirana orodja brez ustreznega človeškega nadzora, ki bi upošteval tudi druge pomembne informacije, bi lahko prispevalo k diskriminaciji pri odločanju.

4.7 PRAVICA DO SOCIALNE VARNOSTI IN SOCIALNE POMOČI

Pravica do socialne varnosti in pomoči iz člena 34 Listine je klasična socialna pravica[63], ki se zgleduje po različnih mednarodnih in evropskih pravnih standardih[64]. Ta določba, ki združuje elemente pravice in načela[65], ima v EU velik pomen zlasti zaradi prostega pretoka ljudi v Uniji.

Namesto da bi vprašanja socialne zaščite povezali s trgom dela, zavzema ta pravica iz Listine nov skupnostni pristop in se sklicuje na zagotavljanje „dostopa do dajatev socialne varnosti in socialnih služb, ki nudijo varstvo v primerih, kot so materinstvo, bolezen, nesreče pri delu, odvisnost ali starost ter v primeru izgube zaposlitve" (člen 34(1))[66].

Gre pa predvsem za programsko izjavo, ki ne predpisuje nobenega minimalnega standarda zaščite. Načeloma je naloga držav članic EU, da določijo pogoje upravičenosti in dostopa do socialnih prejemkov, pri čemer je potrebno tudi dodatno pojasnilo Sodišča[67]. Vendar člen 34(1) Listine zagotavlja zaščito pred ukrepi, ki omejujejo ali odpravljajo obstoječe pravice socialne varnosti[68].

Poleg tega je ob upoštevanju zakonodaje EU in nacionalne zakonodaje dostop do socialnih pravic zagotovljen vsem posameznikom, ki zakonito prebivajo v EU in uveljavljajo pravico do prostega gibanja, ne glede na njihovo državljanstvo (člen 34(2)). Na ta način prihaja do nastanka iztožljivih pravic pred nacionalnimi sodišči in Sodiščem EU[69].

Vse bolj očitno postaja, da je vpliv tehnologij umetne inteligence na sisteme socialne zaščite in življenja številnih posameznikov, ki so odvisni od njih, lahko daljnosežen in potencialno zelo problematičen. Uvajanje tehnologij umetne inteligence v sisteme socialnega varstva lahko predstavlja oviro za dostop do te pravice[70].

Pri uporabi umetne inteligence na področju socialne varnosti je treba na primer upoštevati morebitne negativne in diskriminatorne učinke na nedržavljane (državljane EU in državljane tretjih držav), ki uveljavljajo svojo pravico do prostega gibanja v EU. Do tega bi lahko prišlo denimo, če bi sistem temeljil na podatkih o preteklih zaposlitvah, ki pri tistih, ki so se priselili iz drugih držav članic EU, niso na voljo.

Samo eden od anketirancev je obravnaval „pravico do ustrezne pokojnine" kot enega od vidikov širše definicije človekovih pravic. Nihče od anketiranih pa se v razgovorih ni skliceval na temeljno pravico do socialne varnosti in socialne pomoči. To je lahko delno posledica narave izbranih primerov uporabe. Vendar pa je dejstvo, da zaposleni v javnem sektorju niso omenjali teh pravic, zgovoren podatek.

4.8 VARSTVO POTROŠNIKOV

Listina določa, da morajo politike EU zagotoviti visoko raven varstva potrošnikov, ki temelji na členu 169 PDEU. To načelo morajo upoštevati institucije in drugi organi EU kot tudi organi posameznih držav članic pri izvajanju prava EU[71].

To načelo iz Listine pomeni jamstvo za zagotavljanje določenega cilja („visoke ravni varstva potrošnikov"). Člen 169 PDEU je bolj konkreten, saj določa tudi načine za dosego navedenega cilja, na primer varovanje zdravja, varnosti in ekonomskih interesov potrošnikov, pa tudi spodbujanje njihove pravice do obveščenosti, izobraževanja in samoorganiziranja za zaščito njihovih interesov[72].

Med obravnavanimi primeri uporabe sta še posebej pomembni uporaba umetne inteligence pri ciljno usmerjenem oglaševanju in uporaba zdravstvene dokumentacije s strani zasebnih podjetij.

Pri ciljno usmerjenem oglaševanju se morajo potrošniki zavedati, da imajo možnost zavrniti sodelovanje. Če tega ne vedo, bodo morda prisiljeni dobivati oglase, ki jih ne želijo. To je še posebej problematično v kombinaciji z zelo izpopolnjenimi sistemi umetne inteligence, ki se uporabljajo pri oglaševanju in lahko vodijo do manipulacije s preferencami potrošnikov[73].

Varstvo podatkov je izjemno pomembno tudi pri uporabi zdravstvenih podatkov (EHR). Evropska potrošniška organizacija (BEUC) je ugotovila, da umetna inteligenca na področju zdravstva prinaša številne izzive za potrošnike. Podala je priporočila, da morajo tehnologije umetne inteligence v celoti spoštovati predpise o varstvu podatkov, biti pregledne za potrošnike in preprečevati diskriminacijo. Evropska potrošniška organizacija je prav tako pozvala k posodobitvi uredbe in zakonodajnih ukrepov za nadzor trga, kazenski pregon in učinkovita pravna sredstva v zvezi z izdelki in storitvami na področju digitalnega zdravja z namenom popolne zaščite potrošnikov v EU[74].

V izbranih državah članicah EU je organizacija BEUC med potrošniki izvedla raziskavo o njihovih pogledih na umetno inteligenco. V raziskavi se je izkazalo, da je več kot eden od dveh anketirancev pritrdil dejstvu, da podjetja uporabljajo umetno inteligenco za manipuliranje s potrošniškimi odločitvami. Poleg tega skoraj polovica vprašanih meni, da prilagojene vsebine in oglasi na platformah za e-trgovanje nimajo dodane vrednosti (44 %). Nekaj več kot polovica anketirancev je izrazila nizko stopnjo zaupanja v dejstvo, da vlade učinkovito nadzorujejo uporabo umetne inteligence[75].

V intervjujih, opravljenih za to študijo, je bilo varstvo potrošnikov v zvezi z razpravo o tveganjih uporabe umetne inteligence in temeljnih pravicah omenjeno zgolj obrobno. Vendar nekateri anketiranci iz podjetij omenjajo zakonodajo o varstvu potrošnikov kot ustrezen okvir, ki velja tudi za njihovo uporabo umetne inteligence. Poleg tega nekateri anketiranci menijo, da predstavljajo organi za varstvo potrošnikov potencialno pomemben nadzorni organ pri uporabi umetne inteligence.

Na splošno mnogi anketiranci iz poslovnega sektorja poudarjajo pomen zadovoljstva potrošnikov. Podjetje, ki uporablja videonadzor za varnost strank v svojih prostorih, na primer omenja, da so za take tehnične rešitve pomembni predpisi o varstvu potrošnikov in da bi morala biti uporaba sistemov namenjena izboljšanju položaja potrošnikov in hkrati ohranjanju njihovih pravic. Za razumevanje in profiliranje potrošnikov je bilo oblikovanih več orodij umetne inteligence, ki podjetjem omogočajo, da izboljšajo svoje storitve in trženje.

Varstvo podatkov je pomemben vidik poslovanja. To je povezano tudi z dejstvom, da se kršitev pravil o varstvu podatkov šteje za poslovno tveganje, kot je navedeno zgoraj. Ena od glavnih skrbi podjetij je pridobivanje in upravljanje privolitev potrošnikov in strank za obdelavo njihovih podatkov pri uporabi orodij umetne inteligence za trženje. Anketiranci poročajo, da je splošna uredba o varstvu podatkov vplivala na izboljšanje njihovih sistemov za obravnavo privolitev.

4.9 PRAVICA DO DOBREGA UPRAVLJANJA

Pravica do dobrega upravljanja je uveljavljeno splošno načelo prava EU, ki ga je razvilo Sodišče EU. Zato je zavezujoče za vse države članice EU[76]. Predstavlja tudi temeljno pravico iz člena 41 Listine, čeprav se nanaša le na ukrepe institucij, organov in agencij EU[77].

Kot splošno načelo prava EU zahteva, da države članice EU pri vseh javnih ukrepih uporabljajo zahteve iz naslova pravice do dobrega upravljanja. Ta pravica vključuje, vendar ni omejena na, pravico posameznika, da ima dostop do svojega spisa, in obveznost katerega koli javnega organa, da navede zadostne razloge za svoje odločitve[78].

Dostop do spisa olajša razumevanje dokazne podlage, na kateri je bila sprejeta odločitev, in/ali razlogov, zaradi katerih je bila sprejeta. To posameznika postavlja v boljši položaj za predložitev nasprotnih argumentov pri uveljavljanju pravice do zaslišanja in pravice do učinkovitega pravnega sredstva[79].

Obveznost navedbe razlogov omogoča večjo preglednost postopka odločanja z vidika prizadetih posameznikov, tako da lahko zadevna oseba ve in razume, zakaj je bil ukrep ali postopek sprejet. Preglednost kot načelo prav tako postavlja temelje za druge pravice[80], vključno z uveljavljanjem pravice do učinkovitega pravnega sredstva.

Po mnenju Sodišča je pri določanju obsega dolžnosti navedbe razlogov pomemben kontekst, v katerem se sprejemajo posamezne odločitve[81]. V Franciji na primer velja Kodeks o odnosih med javnostjo in upravo, ki predpisuje pisna pojasnila dejanskih in pravnih vidikov, na katerih temelji odločitev[82].

Pravica do dobrega upravljanja velja tudi takrat, ko obdelava osebnih podatkov poteka s pomočjo umetne inteligence ali ko sistemi umetne inteligence podpirajo odločanje javnih organov. Čeprav je pravica do dobrega upravljanja lahko predmet določenih omejitev, se postavlja vprašanje, kako zagotoviti, da bo imelo potencialno veliko število posameznikov dostop do svojih datotek (osebnih podatkov, ki se uporabljajo v sistemih umetne inteligence). Drugo vprašanje je, kako zagotoviti, da javni organi vedno navedejo zadostne razloge, in to tudi v primerih uporabe tehnologij umetne inteligence, ki jih je zaradi vgrajene kompleksnosti in nejasnosti težko pojasniti.

Uporaba sistema za kategorizacijo brezposelnih oseb, ki je bil vzpostavljen na Poljskem, je poudarila težave, povezane z uporabo algoritmov v javni upravi. Na podlagi vprašanj, na katera so odgovarjali brezposelni, je bila s pomočjo statističnega algoritma razvita posebna kategorizacija. Sistem je bil deležen številnih kritik civilne družbe, zlasti zaradi odsotnosti mehanizma za pritožbe in morebitnega pojava diskriminacije[83]. Na koncu je pritožba varuha človekovih pravic, ki je temeljila na upravnih razlogih, privedla do odločitve ustavnega sodišča, s katero je bila uporaba sistema ustavljena[84].

V javnem sektorju spodbuja uporabo umetne inteligence želja po povečanju učinkovitosti, ki je tesno povezana z izboljšanjem uprave in povečanjem koristi za državljane. Anketiranci s področja javne uprave so kot razlog, zakaj organi razmišljajo o uporabi umetne inteligence ali jo že uporabljajo, najpogosteje navajali ravno učinkovitost. Eden od anketirancev, ki se ukvarja s svetovanjem ministrstvom o digitalnih strategijah in uporabi umetne inteligence, je navedel, da sta glavna razloga za uvajanje umetne inteligence izboljšanje storitev za državljane in zmanjšanje stroškov teh storitev za javno upravo.

Anketiranci so prav tako povedali, da obstajajo v javni upravi posebne zahteve, zaradi katerih umetne inteligence ni mogoče uporabiti za vse namene in zlasti na področju sprejemanja odločitev zahteva posebno pozornost. Vendar pa se učinkovitost sistema prav tako šteje za pomembno dodano vrednost.

V tem smislu je anketiranec, ki se ukvarja z digitalizacijo upravljanja migracij, nakazal, da prinaša oblikovanje preveč zapletenih sistemov umetne inteligence tveganje, saj bi pozneje za razumevanje sistema za nazaj bilo potrebnega veliko dela. Anketiranec navaja, da morajo biti člani njegove ekipe zelo previdni, da umetna inteligenca ne bi sprejemala končnih odločitev, ki jih morajo sprejemati samo ljudje, saj družba in stranke po njegovem mnenju za to niso pripravljene. Čeprav so nekateri sistemi privlačni, njihovo delovanje ni učinkovito, kar bi lahko povzročilo dodatno delo in negativne rezultate. Vendar pa sogovornik prav tako navaja, da je pojem učinkovitosti v razpravah o varstvu podatkov pogosto na stranskem tiru.

Zahteve za dobro upravljanje prav tako neposredno povezujejo zgoraj navedena vprašanja z vidiki varstva podatkov, varstva pred diskriminacijo ter pravice do učinkovitega pravnega sredstva in poštenega sojenja. Javna uprava lahko obdeluje podatke le na ustrezni pravni podlagi. Odločitve morajo biti poštene in pregledne, poti za izpodbijanje odločitev pa morajo biti razpoložljive in dostopne. Zato so zahteve za dobro upravljanje neposredno povezane z zgoraj navedeno razpravo in analizo v zvezi z zakonito obdelavo podatkov (v okviru varstva podatkov), sprejemanjem pravičnih odločitev (v okviru razprave o varstvu pred diskriminacijo) ter preglednostjo in možnostmi izpodbijanja in obrazložitve odločitev (v okviru dostopa do pravnega varstva).

Končne opombe

1 Evropska komisija, *European enterprise survey on the use of technologies based on artificial intelligence* (Raziskava evropskih podjetij o uporabi tehnologij, ki temeljijo na umetni inteligenci), Luxembourg, julij 2020.
2 FRA (2020), ***Kaj temeljne pravice pomenijo ljudem v EU?***, Luxembourg, Urad za publikacije, str. 28.
3 Glej **spletno stran o tristopenjskem informacijskem varnostnem sistemu ISKE** na spletnem mestu estonskega organa za informacijske sisteme.
4 Glej **spletno mesto Odbora za varnostne standarde PCI**.
5 Barak, A. (2019), *Human dignity as a framework right (motherright)* (Človekovo dostojanstvo kot okvirna pravica), v Barak, A., *Human Dignity: The Constitutional Value and the Constitutional Right* (Človekovo dostojanstvo: ustavna vrednota in ustavna pravica), Cambridge, Cambridge University Press, 2015, poglavje 9, str. 156–169.
6 Sodba Sodišča z dne 9. oktobra 2001, Nizozemska proti Evropskemu parlamentu in Svetu, C-377/98, točke 70–77.
7 Za razpravo o zlonamerni uporabi umetne inteligence glej npr. Brundage, M. et al. (2018), ***The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*** (Zlonamerna uporaba umetne inteligence: napovedovanje, preprečevanje in ublažitev).
8 FRA (2019), ***Facial recognition technology: fundamental rights considerations in the context of law enforcement*** (Tehnologija za prepoznavanje obraza: temeljni vidiki pravic v okviru kazenskega pregona), Luxembourg, Urad za publikacije, november 2019.
9 Sklepni predlogi generalne pravobranilke Sharpston z dne 17. junija 2010, Volker und Markus Schecke GbR in Hartmut Eifert proti Land Hessen, združeni zadevi C-92/09 in C-93/09, točka 71.
10 FRA, Svet Evrope in ENVP (2018), ***Priročnik o evropskem protidiskriminacijskem pravu. Izdaja iz leta 2018***, Luxembourg, Urad za publikacije, junij 2018, str. 19.
11 Glej tudi prav tam, str. 35–52.
12 ESČP (2019), *Guide on Article 8 of the European Convention on Human Rights – Right to respect for private and family life, home and correspondence* (Priročnik za razlago 8. člena EKČP – pravica do spoštovanja zasebnega in družinskega življenja, stanovanja ter komunikacij), Strasbourg, Svet Evrope, posodobljeno 31. avgusta 2019, odst. 133 in 136.
13 ESČP, López Ribalda in drugi proti Španiji, št. 1874/13 in 8567/13, 17. oktober 2019, točka 87. Za celovito pravno analizo pomena in vsebine „zasebnosti“ glej tudi Koops, B.-J. et al. (2017), „**A Typology of Privacy**“ (Tipologija zasebnosti), *University of Pennsylvania Journal of International Law*, zvezek 38, izdaja 2, str. 483–575.
14 Vermeulen, M. (2015), ***SURVEILLE Deliverable D4.7 – The scope of the right to private life in public places*** (SURVEILLE, rezultat 4.7 – obseg pravice do zasebnega življenja na javnih mestih), julij 2014, str. 2.
15 OZN, Odbor za človekove pravice, ***Splošni komentar št. 37 (2020) o pravici do mirnega zbiranja (21. člen)***, CCPR/C/GC/37, 17. september 2020, str. 62.
16 Costello, R. Á. (2020), „**The Impacts of AdTech on Privacy Rights and the Rule of Law**“ (Učinki oglaševalske tehnologije na pravice do zasebnosti in vladavino prava), Technology and Regulation.
17 Norveški potrošniški svet (2020), *Out of Control. How consumers are exploited by the online advertising industry* (Brez nadzora, kako spletna oglaševalska industrija izkorišča potrošnike).
18 FRA (2020), ***Your rights matter: Data protection and privacy – Fundamental Rights Survey*** (Vaše pravice so pomembne: varstvo podatkov in zasebnosti – pregled temeljnih pravic), Luxembourg, Urad za publikacije.
19 Rocher, L., Hendrickx, J. M. in de Montjoye, Y. (2019), „**Estimating the success of re-identifications in incomplete datasets using generative models**“ (Ocena uspešnosti ponovne identifikacije v nepopolnih zbirkah podatkov z uporabo generativnih modelov), *Nature Communications* 10, št. 3069.
20 Hacker, P. (2020), *A Legal Framework for AI Training Data. Law, Innovation and Technology* (Pravni okvir učnih podatkov umetne inteligence. Pravo, inovacije in tehnologija, v pripravi), na voljo na **SSRN**; Delovna skupina za varstvo podatkov iz člena 29 (2014), ***Mnenje št. 5/2014 o anonimizacijskih tehnikah***; glej tudi Finck, M. in Pallas, F., *They Who Must Not Be Identified - Distinguishing Personal from Non-Personal Data Under the GDPR* (Tisti, ki jih ni dovoljeno identificirati – ločevanje osebnih od neosebnih podatkov v skladu z GDPR, 1. oktober 2019), v pripravi, *International Data Privacy Law* (Mednarodno pravo o varstvu podatkov), 2020, Inštitut Maxa Plancka za inovacije in konkurenco, raziskovalni dokument št. 19–14), na voljo na **SSRN**; ter Sartor, G. in Lagioia, F. (2020), *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence* (Vpliv splošne uredbe o varstvu podatkov na umetno inteligenco), študija, pripravljena za Odbor za prihodnost znanosti in tehnologije (STOA) Evropskega parlamenta.
21 Glej npr. blog Službe Združenega Kraljestva za podatke „**Access to sensitive data for research: ‘The 5 Safes’**“ (Dostop do občutljivih podatkov za raziskovalne namene: pet varoval); glej tudi razpravo v Ohm, P. (2010), „Broken promises of privacy: responding to the surprising failure of anonymization“ (Neizpolnjene obljube zasebnosti: odziv na presenetljiv neuspeh anonimizacije), *UCLA Law Review*, str. 1701.
22 GDPR, člen 22(3), in direktiva o kazenskem pregonu, člen 11(1).
23 Veale, M. in Edwards, L. (2018), „Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling“ (Preglednost, presenečenja in dodatna vprašanja v zvezi s smernicami za avtomatizirano odločanje in profiliranje delovne skupine iz člena 29), *Computer Law & Security Review*, zvezek 34(2), april 2018, str. 398–404.
24 Delovna skupina za varstvo podatkov iz člena 29 (2018), ***Smernice o avtomatiziranem sprejemanju posameznih odločitev in oblikovanju profilov za namene Uredbe (EU) 2016/679***, sprejete 3. oktobra 2017, kot so bile nazadnje revidirane in sprejete 6. februarja 2018.
25 Green, B. in Chen, Y. (2019), „Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments“, (Disparitetne interakcije: analiza algoritma v zanki z vidika poštenosti in ocene tveganja), v *FAT* '19: Conference on Fairness, Accountability, and Transparency* (FAT* '19: Konferenca o poštenosti, odgovornosti in transparentnosti), 29.–31. januar 2019.
26 González Fuster, G. (2020), ***Artificial Intelligence and Law Enforcement – Impact on Fundamental Rights*** (Umetna inteligenca in kazenski pregon – učinek na temeljne pravice), Evropski parlament, Tematski sektor za pravice državljanov in ustavne zadeve, Generalni direktorat za notranjo politiko, PE 656.295, julij 2020, str. 17; Brkan, M. (2019), „Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond“ (Ali algoritmi vladajo svetu? Algoritemsko odločanje in varstvo podatkov v okviru GDPR ter širše), *International Journal of Law and Information Technology*, zvezek 27 (2), str. 98; Delovna skupina iz člena 29, ***Smernice o avtomatiziranem sprejemanju posameznih odločitev in oblikovanju profiliranju za namene Uredbe (EU) 2016/679***, sprejete 3. oktobra 2017, kot so bile nazadnje revidirane in sprejete 6. februarja 2018, WP251rev.0, str. 19.
27 Misuraca, G. in van Noordt, C. (2020), ***Overview of the use and impact of AI in public services in the EU*** (Pregled uporabe in učinkov umetne inteligence na področju javnih storitev v EU), Skupno raziskovalno središče Evropske komisije, Luxembourg.
28 Direktiva Sveta 2000/43/ES z dne 29. junija 2000 o izvajanju načela enakega obravnavanja oseb ne glede na raso ali narodnost, UL L 180, 19.7.2000, str. 22–26, člen 2, in Direktiva Sveta 2000/78/ES z dne 27. novembra 2000 o splošnih okvirih enakega obravnavanja pri zaposlovanju in delu, UL L 303, 2.12.2000, str. 16–22, člen 2.

29 FRA in Svet Evrope (2018), ***Priročnik o evropskem protidiskriminacijskem pravu. Izdaja iz leta 2018***, Luxembourg, Urad za publikacije, junij 2018, str. 35.
30 Sodba Sodišča z dne 1. marca 2011, Association Belge des Consommateurs Test-Achats ASBL in drugi proti Conseil des ministres, C-236/09.
31 Evropska komisija (2012), V Evropski uniji začnejo veljati predpisi o oblikovanju cen v zavarovalniški industriji brez razlikovanja glede na spol, sporočilo za javnost, **IP/11/1581**, 20. december 2012.
32 Joh, E. E. (2015), „**The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing**“ (Nova diskretnost nadzora: avtomatizirani sum, masovni podatki in nadzor), *UC Davis Legal Studies Research Paper No. 473*, str. 17–18.
33 Završnik, A. (2019), „Algorithmic justice: Algorithms and big data in criminal justice settings“ (Algoritemsko pravosodje: Algoritmi in masovni podatki v kazenskem pravosodju), *European Journal of Criminology*, str. 14. DOI: 10.1177/1477370819876762.
34 Glej tudi Evropska komisija, *Bela knjiga o umetni inteligenci – evropski pristop k odličnosti in zaupanju*, COM(2020) 65 final, Bruselj, 19. februar 2020, str. 1.
35 FRA (2018), ***#BigData: Discrimination in data-supported decision making*** (#Masovni podatki: Diskriminacija pri sprejemanju odločitev, podprtih s podatki), Luxembourg, Urad za publikacije, junij 2018, str. 3.
36 Prav tam.
37 FRA (2019), ***Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*** (Kakovost podatkov in umetna inteligenca – blažitev pristranskosti in napak pri varstvu temeljnih pravic), Luxembourg, Urad za publikacije.
38 Korff, D. in Browne, I. (2013) „**The use of the Internet & related services, private life & data protection: trends, technologies, threats and implications**“ (Uporaba interneta in sorodnih storitev, zasebno življenje in varstvo podatkov: trendi, tehnologije, grožnje in posledice), Svet Evrope, T-PD(2013)07.
39 Glej Nacionalno sodišče za varstvo pred diskriminacijo in enakost Finske, **sklep št. 216/2017** z dne 21. marca 2018. Glej tudi zgoraj obravnavano zadevo SyRI in Združeno kraljestvo, Pritožbeno sodišče, **R (Bridges) proti CC South Wales**, [2020] EWCA Civ 1058, 11. avgust 2020.
40 Glej tudi Equinet (2020), ***Regulating for an equal AI: A new role for equality bodies*** (Pravno urejanje za enakopravno umetno inteligenco: nova vloga za organe za enakost), Bruselj, poročilo, ki sta ga pripravila Allen, R. in Masters, D.
41 Tolan, S., Miron, M., Gomez, E. in Castillo, C. (2019), „Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia“ (Zakaj lahko strojno učenje privede do neenakosti: dokazi na podlagi ocene tveganja pri mladoletniškem prestopništvu v Kataloniji), nagrada za najboljši članek, Mednarodna konferenca o umetni inteligenci in pravu, 2019; Richardson, R., Schultz, J. in Crawford K. (2019), „Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice“ (Umazani podatki, slabe napovedi: Kako kršitve državljanskih pravic vplivajo na policijske podatke, sisteme napovednega policijskega dela in pravičnost), *94 N. Y. U. L. REV. ONLINE 192 (2019)*, na voljo na **SSRN**.
42 FRA (2014), ***Violence against women: an EU-wide survey. Main results report*** (Nasilje nad ženskami: vseevropska raziskava. Poročilo o glavnih rezultatih), Luxembourg, Urad za publikacije, str. 61.
43 FRA (2018), ***Second European Union Minorities and Discrimination Survey. Main results*** (Druga raziskava o manjšinah in diskriminaciji v Evropski uniji. Glavni rezultati), Luxembourg, Urad za publikacije, str. 66.
44 Bakke, E. (2018), „Predictive policing: The argument for public transparency“ (Napovedno policijsko delo: argument v prid javni preglednosti), *New York University Annual Survey of American Law*, zvezek 74, str. 139–140; Ferguson, A. G. (2017), „**Policing Predictive Policing**“ (Nadzor napovednega policijskega dela), *Washington University Law Review*, zvezek 94, str. 1146–1150; in Odbor strokovnjakov za internetne posrednike Sveta Evrope (MSI-NET) (2017), ***Algorithms and Human Rights*** (Algoritmi in človekove pravice), Svet Evrope DGI(2017)12, str. 11.
45 Joh, E. E. (2015), „**The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing**“ (Nova diskretnost nadzora: avtomatizirani sum, masovni podatki in nadzor), *UC Davis Legal Studies Research Paper No. 473*, str. 18.
46 Gstrein, O. J., Bunnik, A. in Zwitter, A. (2019), „Ethical, Legal and Social Challenges of Predictive Policing“ (Etični, pravni in družbeni izzivi napovednega policijskega dela), *Católica Law Review* 3:3, str. 77–98; Meijer, A. in Wessels, M. (2019), „Predictive Policing: Review of Benefits and Drawbacks“ (Napovedno policijsko delo: pregled koristi in pomanjkljivosti), *International Journal of Public Administration* 42:12, str. 1036, DOI: 10.1080/01900692.2019.1575664.
47 Završnik, A. (2019), „Algorithmic justice: Algorithms and big data in criminal justice settings“ (Algoritemsko pravosodje: Algoritmi in masovni podatki v kazenskem pravosodju), *European Journal of Criminology*, str. 8–9. DOI: 10.1177/1477370819876762.
48 Wachter, S. (2020), „Affinity Profiling and Discrimination by Association in Online Behavioural Advertising“ (Afinitativno profiliranje in diskriminacija zaradi pripadnosti v spletnem vedenjskem oglaševanju), *Berkeley Technology Law Journal*, zvezek 35, št. 2, 2020, (v pripravi), na voljo na **SSRN**.
49 Za več informacij o uporabi umetne inteligence v finančnem sektorju, ki vodi do neenakega dostopa do finančnih storitev, glej v pravni literaturi npr. Boyd, D., Levy, K. in Marwick, A. (2014), „The Networked Nature of Algorithmic Discrimination“ (Mrežna narava algoritemske diskriminacije), v Gangadharan, S. P., Eubanks, V. in Barocas, S. (ur.), ***Data and Discrimination: Collected Essays*** (Podatki in diskriminacija: zbrani eseji), Open Technology Institute, str. 53–62.
50 Za pregled vprašanj, povezanih z otrokovimi pravicami, glej UNICEF, Center za inovacije in človekove pravice, UC Berkeley (2019), ***Artificial Intelligence and Children's Rights*** (Umetna inteligenca in otrokove pravice).
51 Mreža neodvisnih strokovnjakov EU za vprašanja temeljnih pravic, ***Commentary on the Charter on Fundamental Rights of the European Union*** (Komentar Listine Evropske unije o temeljnih pravicah), junij 2006, str. 360. Glej tudi: FRA in Svet Evrope (2016), ***Priročnik o evropski zakonodaji v zvezi z dostopom do pravnega varstva***, Luxembourg, Urad za publikacije, junij 2016, str. 92.
52 Sodba Sodišča z dne 13. marca 2007, Unibet (London) Ltd in Unibet (International) Ltd proti Justitiekanslern, C-432/05, točka 37; sodba Sodišča z dne 27. junija 2013, ET Agrokonsulting-04-Velko Stojanov proti Izpalnitelen direktor na Daržaven fond „Zemedelie“ – Razplaštatelna agencija, C-93/12, točka 59; sodba Sodišča z dne 18. decembra 2014, Centre public d'action sociale d' Ottignies-Louvain-la-Neuve proti Moussi Abdidi, C-562/13, točka 45.
53 Direktiva o kazenskem pregonu, člen 54, in GDPR, člen 79.
54 Direktiva o kazenskem pregonu, člen 53, in GDPR, člen 78.
55 Direktiva o kazenskem pregonu, člen 52, in GDPR, člen 77.
56 Svet Evrope, ***Priporočilo CM/Rec(2020)1 Odbora ministrov državam članicam o posledicah algoritemskih sistemov za človekove pravice*** (ki ga je Odbor ministrov sprejel 8. aprila 2020 na 1373. seji namestnikov ministrov), dodatek, odst. B.4.5.
57 Ferguson, A. G. (2017), „**Policing Predictive Policing**“ (Nadzor napovednega policijskega dela), 94 *Washington University Law Review*, str. 1165–1167.
58 Gstrein, O. J., Bunnik, A. in Zwitter, A. (2019), „Ethical, Legal and Social Challenges of Predictive Policing“ (Etični, pravni in družbeni izzivi napovednega policijskega dela), *Católica Law Review*, 3:3, str. 80–81; Yeung, K. (2019), ***A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework*** (Študija posledic naprednih digitalnih tehnologij (vključno s sistemi umetne inteligence) za koncept odgovornosti v okviru človekovih pravic), ki jo je pripravil Odbor

strokovnjakov Sveta Evrope za vidike človekovih pravic pri avtomatski obdelavi podatkov in različnih oblikah umetne inteligence (MSI-AUT).

59 Svet Evrope, *Algorithms and human rights* (Algoritmi in človekove pravice), str. 11 in 24.

60 Združenje ITechLaw (2019), ***Responsible AI: A Global Policy Framework*** (Odgovorna umetna inteligenca: globalni okvir politike), str. 258–282.

61 Pomanjkanje strokovnega znanja o umetni inteligenci se je pokazalo tudi v raziskavi podjetij v EU, pomanjkanje znanja med obstoječim osebjem in težave pri zaposlovanju novega osebja namreč predstavljajo najpomembnejše ovire za nadaljnje uvajanje umetne inteligence (Evropska komisija, (2020), *European enterprise survey on the use of technologies based on artificial intelligence* (Raziskava evropskih podjetij o uporabi tehnologij, ki temeljijo na umetni inteligenci), Luxembourg, julij 2020, str. 11).

62 Glej podrobno oceno učinkov napovednega policijskega dela na domnevo nedolžnosti v Mendola, M. (2016), ***One Step Further in the 'Surveillance Society': The Case of Predictive Policing*** (Korak naprej k družbi nadzorovanja: primer napovednega policijskega dela).

63 Glej npr. Egorov, A. in Wujczyk, M. (ur.) (2016), ***The Right to Social Security in the Constitutions of the World: Broadening the moral and legal space for social justice*** (Pravica do socialnega varstva v ustavah sveta: širitev moralnega in pravnega prostora za socialno pravičnost), Ženeva, globalna študija MOD, zvezek 1: Evropa, str. xv–xvii in 1–6.

64 Ti vključujejo čl. 153 in 156 PDEU; čl. 12 in 13 Evropske socialne listine iz leta 1961 ter točki 2 in 10 Listine Skupnosti o temeljnih socialnih pravicah delavcev iz leta 1989 (glej **Pojasnila k Listini o temeljnih pravicah**, UL C 303, 14.12.2007, str. 17–35).

65 **Pojasnila k Listini o temeljnih pravicah** (UL C 303, 14.12.2007, str. 17–35), Pojasnilo k členu 52 – Obseg pravic in načel ter njihova razlaga.

66 Bojarski, Ł., Schindlauer, D. in Wladasch, W. (2014), *The European Charter of Fundamental Rights as a Living Instrument – Manual* (Listina Evropske unije o temeljnih pravicah kot živi instrument – priročnik), Rim/Varšava/Dunaj, str. 61–62.

67 De Becker, E. (2016), „**The (Possible) Role of the Right to Social Security in the EU Economic Monitoring Process**" ((Morebitna) vloga pravice do socialne varnosti v procesu ekonomskega monitoringa v EU), *German Law Journal*, zvezek 17, št. 3, str. 297, 304; Paju, J. (2017), *The European Union and Social Security Law* (Evropska unija in pravo socialne varnosti), Oxford, Hart Publishing, pododdelek 7.5.2.

68 Prav tam, str. 297–298; Peers, S. in Prechal, S. (2014), „Scope and Interpretation of Rights and Principles" (Obseg in razlaga pravic in načel), v Hervey, T., Kenner, J., Peers, S. in Ward, A. (ur.), *The EU Charter of Fundamental Rights. A Commentary* (Listina EU o temeljnih pravicah. Komentar), Oxford in Portland, Oregon; Hart Publishing, 2014, str. 1455, 1508.

69 Z izjemo Poljske in Združenega kraljestva, glej Protokol (št. 30) o uporabi Listine o temeljnih pravicah Evropske unije na Poljskem in v Združenem kraljestvu (UL C 115, 9.5.2008, str. 313–314), člen 1(2).

70 van Veen, C. in Zevenbergen, B. „**Conference on Social Protection by Artificial Intelligence: Decoding Human Rights in a Digital Age**" (Konferenca o socialni zaščiti z uporabo umetne inteligence: dešifriranje človekovih pravic v digitalni dobi), *Freedom to Tinker – Research and Expert Commentary on Digital Technologies in Public Life*, 29. maj 2019.

71 Čl. 51(1) Listine; glej tudi **Pojasnila k Listini o temeljnih pravicah** (UL C 303, 14.12.2007, str. 17–35), Pojasnilo k členu 52 – Obseg pravic in načel ter njihova razlaga.

72 Bojarski, Ł., Schindlauer, D. in Wladasch, K. (2014), *The European Charter of Fundamental Rights as a Living Instrument – Manual* (Listina Evropske unije o temeljnih pravicah kot živi instrument – priročnik), Rim/Varšava/Dunaj, str. 67.

73 Sartor, G. (2020), ***New aspects and challenges in consumer protection*** (Novi vidiki in izzivi varstva podatkov), študija za Odbor za notranji trg in varstvo potrošnikov, Tematski sektor za gospodarsko in znanstveno politiko ter kakovost življenja, Evropski parlament, Luxembourg.

74 Evropska potrošniška organizacija (BEUC) (2018), ***Digital Health, Principles and Recommendations*** (Digitalno zdravje, načela in priporočila).

75 BEUC (2020), ***Artificial Intelligence: what consumers say. Findings and policy recommendations of a multi-country survey on AI*** (Umetna inteligenca: kaj pravijo potrošniki. Ugotovitve in priporočila glede politike v okviru meddržavne raziskave o umetni inteligenci).

76 V nedavni sodni praksi glej sodbo Sodišča z dne 8. maja 2014, H. N. proti Minister for Justice, Equality and Law Reform, Ireland, Attorney General, C-604/12, točka 49.

77 Potrjeno tudi v sodbi Sodišča z dne 17. julija 2014, YS proti Minister voor Immigratie, Integratie en Asiel in Minister voor Immigratie, Integratie en Asiel proti M, S, združeni zadevi C-141/12 in C-372/12, točke 66–70.

78 Ti elementi, ki jih je v svoji sodni praksi najprej razvilo Sodišče EU, so zapisani v členu 41(2) Listine. Za več informacij o tej pravici v vodilni akademski literaturi glej Craig, P. (2014), „Article 41 – Right to Good Administration" (Člen 41 – pravica do dobrega upravljanja), v Hervey, T., Kenner, J., Peers, S. in Ward, A. (ur.), *The EU Charter of Fundamental Rights. A Commentary* (Listina EU o temeljnih pravicah. Komentar), Oxford in Portland, Oregon; Hart Publishing, 2014, str. 1069–1098.

79 Prav tam, str. 1082.

80 Finck, M. (2019), „**Automated Decision-Making and Administrative Law**" (Avtomatizirano odločanje in upravno pravo), *Inštitut Maxa Plancka za inovacije in konkurenco: raziskovalni dokument št. 19–10*, str. 8.

81 Craig, P. (2014), „Article 41 – Right to Good Administration" (Člen 41 – pravica do dobrega upravljanja), v Hervey, T., Kenner, J., Peers, S. in Ward, A. (ur.), *The EU Charter of Fundamental Rights. A Commentary* (Listina EU o temeljnih pravicah. Komentar), Oxford in Portland, Oregon; Hart Publishing, 2014, str. 1086–1087.

82 Francija, ***Code des relations entre le public et l'administration***, člen L2111-5.

83 Panoptykon Foundation (2015), ***Profiling the unemployed in Poland: Social and political implications of algorithmic decision making*** (Profiliranje brezposelnih na Poljskem: socialne in politične posledice algoritemskega odločanja); glej tudi Algorithm Watch (2019), ***Poland to scrap controversial unemployment scoring system*** (Poljska umika sporni sistem za ocenjevanje brezposelnosti).

84 Glej **Sklep K 53/16,** ki je na voljo na spletni strani Ustavnega sodišča.

5.

OCENA UČINKA NA TEMELJNE PRAVICE – PRAKTIČNO ORODJE ZA VARSTVO TEMELJNIH PRAVIC

V poglavju 4 smo ponazorili, v kolikšni meri uporaba umetne inteligence vpliva na različne temeljne pravice. To poglavje pa se posveča analizi, kako bi lahko ocena učinka v zvezi s temeljnimi pravicami (angl. *fundamental rights impact assessment*, FRIA) zmanjšala morebitne negativne učinke v zvezi z uporabo umetne inteligence na temeljne pravice.

Oddelek 5.1 vsebuje kratek pregled trenutne razprave o potrebi po oceni učinka v zvezi s temeljnimi pravicami na tem področju. **Oddelek 5.2** vsebuje analizo sedanje prakse pri obravnavanju posledic za temeljne prakse, ki je bila pripravljena na podlagi razgovorov, opravljenih v okviru tega poročila. Anketiranci so bili vprašani o tem, kakšno testiranje je bilo opravljeno pred uporabo sistema in kdo nadzoruje naloge, na katere vpliva uporaba te tehnologije.

Poglavje se konča s predlogi, kako poskrbeti za temeljne pravice pri uporabi umetne inteligence in sorodnih tehnologij.

5.1 POZIV K OCENI UČINKA V ZVEZI S TEMELJNIMI PRAVICAMI – RAZPOLOŽLJIVE SMERNICE IN ORODJA

Mednarodne organizacije[1], akademiki[2] in civilna družba[3] so pozvali k izvedbi ocene učinka umetne inteligence in sorodnih tehnologij na temeljne pravice.

V Smernicah odbora ministrov Sveta Evrope o vplivu algoritemskih sistemov na človekove pravice je na primer podano priporočilo, naj države „pred postopki javnih naročil, med razvojem in ob priložnosti rednih mejnikov pri uvajanju teh sistemov na posameznih področjih izvedejo oceno učinka z namenom prepoznave rezultatov, ki bi lahko bili škodljivi za človekove pravice“[4].

Potrebne so fleksibilne ocene učinka, ki jih je mogoče prilagoditi različnim razmeram, saj so kršitve temeljnih pravic vedno vezane na konkreten kontekst. Strokovnjaki situacijo ponazarjajo z razmerami na področju protidiskriminacijske zakonodaje EU, saj je tudi enakost vedno kontekstualna in odvisna od obravnavanega primera[5].

Spoštovanja temeljnih pravic ni mogoče avtomatizirati ali fiksno kodirati v računalniško programsko opremo. Namesto tega je treba posebej preučiti vsak primer uporabe, da se ugotovi, ali se pojavljajo kakršna koli vprašanja v zvezi s temeljnimi pravicami. Vendar je pri takšnih ocenah mogoče slediti sistematičnemu pristopu, ki zagotavlja podobne informacije.

Obstoječi standardi zagotavljajo smernice, kako izvesti oceno učinka umetne inteligence in sorodnih tehnologij na temeljne pravice. Te vključujejo zavezujoče predpise, instrumente mehkega prava (npr. priporočila ali deklaracije) in praktična orodja (npr. smernice in kontrolne sezname).

Poleg zahtev, ki izhajajo iz zakonodaje s področja varstva podatkov (glej okvir), najdemo le malo primerov zakonov, ki bi na splošno predpisovali obvezno izvedbo ocene učinka umetne inteligence. Zaradi vse intenzivnejše uporabe umetne inteligence je kanadska vlada izdala smernice, vključno z obveznimi zahtevami, ki določajo, kako pravilno pripraviti oceno uporabe umetne inteligence v javni upravi. Uporablja se za vse sisteme, orodja in statistične modele, ki so v uporabi v smislu priporočanja ali sprejemanja upravnih odločb, ki zadevajo stranke[6].

Kaj se lahko naučimo iz ocen učinka v zvezi z varstvom podatkov

Evropska zakonodaja na področju varstva podatkov zahteva izvedbo ocene učinka v zvezi z varstvom podatkov (angl. *data protection impact assessment*, DPIA)[a]. Posodobljena konvencija Sveta Evrope št. 108 določa splošno obveznost, da se pred začetkom obdelave osebnih podatkov preuči verjeten vpliv teh dejavnosti na pravice in temeljne svoboščine posameznikov. Po izvedbi ocene morajo upravljavci zasnovati obdelavo tako, da preprečijo ali vsaj zmanjšajo ugotovljena tveganja[b].

Zakonodaja EU nalaga podobno, a še podrobneje opredeljeno obveznost. GDPR predvideva oceno učinka v zvezi z varstvom podatkov, kadar je možno, „da bi lahko vrsta obdelave [...] povzročila veliko tveganje za pravice in svoboščine posameznikov“[c]. Tako bi morda z oceno učinka umetne tehnologije na področjih, kjer bi jo zakon predpisoval, lahko naslovili njene širše posledice za človekove pravice skupaj z učinki na pravico do zasebnosti[d] in uporabili oceno učinka kot orodje za dodatno raziskovanje algoritmov in njihovih učinkov[e].

Vendar je v skladu z GDPR (člen 35) ocena učinka v zvezi z varstvom podatkov omejena na visoko tvegane primere obdelave osebnih podatkov. Na ta način lahko spregledamo druge visoko tvegane primere, ki niso v prvi vrsti ali na prvi pogled povezani z varstvom osebnih podatkov. Obenem je uporaba splošne uredbe o varstvu osebnih podatkov omejena na zadevno konkretno področje uporabe in povezano strokovno znanje s tega področja. To pomeni, da bi lahko bila možnost razširitve področja uporabe ocene učinka v zvezi z varstvom podatkov na druge temeljne pravice omejena.

GDPR vsebuje tudi nekaj navodil glede načina izvajanja ocene učinka v zvezi z varstvom podatkov. Prvič, ocena učinka bi morala biti izvedena pred kakršno koli visoko tvegano obdelavo podatkov[f]. Drugič, ocena učinka v zvezi z varstvom podatkov bi morala zagotoviti sistematičen opis predvidenih dejanj obdelave, namena in zasledovanih zakonitih interesov. Poleg tega je treba oceniti tudi potrebnost in sorazmernost obdelave ter morebitna tveganja za pravice posameznikov. Ocena učinka mora vsebovati načrtovane varnostne ukrepe za obravnavo ugotovljenih tveganj[g].

Medtem ko poudarja, da se lahko uporabijo različne metodologije, predlaga Delovna skupina iz člena 29 minimalna merila v obliki kontrolnega seznama, ki bi ga upravljalec lahko uporabil za preverjanje, ali ocena učinka v zvezi z varstvom podatkov v celoti ustreza določbam iz splošne uredbe o varstvu podatkov[h].

Nazadnje, GDPR predvideva tudi predhodno obvezno posvetovanje z ustreznim nadzornim organom, če ocena učinka pokaže, da obdelava prinaša tveganja, ki jih ni mogoče ublažiti[i]. To daje organom za varstvo podatkov kot neodvisnim organom, ustanovljenim na podlagi zakona, ključno vlogo[j].

Evropski nadzornik za varstvo podatkov (ENVP) zagotavlja navodila za izvajanje ocen učinka v zvezi z varstvom pravic[k]. Organi za varstvo podatkov so prav tako razpravljali o ocenjevanju tehnologij umetne inteligence in zagotavljajo navodila o tem[l].

[a] *Za več informacij o oceni učinka v zvezi z varstvom podatkov glej: FRA, Svet Evrope in ENVP (2018),* **Priročnik o evropskem pravu varstva podatkov, izdaja iz leta 2018***, str. 179–181.*
[b] ***Posodobljena konvencija Sveta Evrope št. 108****, člen 10(2).*
[c] *GDPR, člen 35(1).*
[d] *GDPR, uvodni izjavi 2 in 75; Delovna skupina iz člena 29, Smernice glede ocene učinka v zvezi z varstvom podatkov, wp248rev.01, 13. oktober 2017.*
[e] *Edwards in Veale (2018); FRA (2018),* **#BigData: Discrimination in data-supported decision making** *(#Masovni podatki: Diskriminacija pri sprejemanju odločitev, podprtih s podatki), Luxembourg, Urad za publikacije, junij 2018.*
[f] *GDPR, člen 35(1). Delovna skupina iz člena 29 poudarja, da je „[i]zvajanje ocene učinka v zvezi z varstvom podatkov [...] stalen proces, ne enkraten dogodek“.*
[g] *GDPR, člen 35(7) ter uvodni izjavi 84 in 90.*
[h] *Delovna skupina iz člena 29,* ***Smernice glede ocene učinka v zvezi z varstvom podatkov****, wp248rev.01, 13. oktober 2017, Priloga 2.*
[i] *GDPR, člen 36.*
[j] *GDPR, člen 35.*
[k] *ENVP (2019),* **Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation** *(Odgovornost na terenu, del II: Ocena učinka v zvezi z varstvom podatkov in predhodno posvetovanje), v1.3, julij 2019.*
[l] *Glej na primer* ***Deklaracijo o etiki in varstvu osebnih podatkov pri umetni inteligenci****, sprejeto leta 2018 na 40. mednarodni konferenci pooblaščencev za varstvo podatkov in zasebnost (ICDPPC).*

Obstaja tudi veliko primerov nezavezujočih smernic. Na svetovni ravni vodilna načela Združenih narodov o podjetništvu in človekovih pravicah priporočajo, da podjetja vključijo ugotovitve iz ocen učinka na človekove pravice v ustrezne notranje funkcije in procese ter ustrezno ukrepajo[7]. Čeprav se ne nanašajo posebej na umetno inteligenco, so smernice pomembne v smislu podpore razvoja umetne inteligence na način, ki bo skladen s pravicami[8].

Na ravni EU velja omeniti Etične smernice za zaupanja vredno umetno inteligenco Strokovne skupine Evropske komisije na visoki ravni za umetno inteligenco[9], ki vsebujejo priporočila za izvajanje ocen učinka na človekove pravice še pred razvojem sistemov, kadar „obstaja tveganje, da bi tehnologija negativno vplivala na temeljne pravice"[10]. Poudarjajo tudi potrebo po vzpostavitvi mehanizmov za prejemanje zunanjih povratnih informacij o sistemih umetne inteligence, ki bi lahko kršili temeljne pravice.

Tudi zasebna podjetja[11], združenja zasebnih podjetij[12] ali javnih in zasebnih akterjev[13] ter nevladne organizacije[14] in druge organizacije[15] so razvili različne vrste smernic za podporo pri izvajanju ocene učinka v zvezi z umetno inteligenco. Vendar ti dokumenti praviloma ne vsebujejo jasnih smernic za izvajanje ocene učinka. Namesto tega poudarjajo različne vidike in merila, ki jih je treba upoštevati pri razvoju in izvedbi ocene učinka.

Široke kategorije vključujejo namen sistema, opis tehnologije, oceno učinka in ciljne populacije/posameznika, vrednotenje z vidika pravičnosti in raznolikosti, opis načrtovanih ali izvedenih revizij ter odgovornost. Nekatere se izrecno sklicujejo na veljavne standarde mednarodnega prava človekovih pravic[16].

Sprejeti so bili tudi različni etični kodeksi ali kodeksi ravnanja[17], standardi[18] in certifikacijske sheme[19].

Na voljo je več praktičnih orodij za oceno učinka tehnologij umetne inteligence in zmanjšanje tveganj, ki so jih razvili številni različni akterji. Mednje spadajo kontrolni seznami[20], seznami vprašanj[21], spletna orodja za samoocenjevanje[22] in okviri za upravljanje s tveganji[23].

Nekatera orodja se osredotočajo posebej na ocenjevanje tveganja na področju temeljnih pravic[24]. Druga se osredotočajo na etične, družbene ali gospodarske posledice[25]. Ta lahko predstavljajo koristen vir podatkov pri izvajanju temeljite ocene učinka tehnologij umetne inteligence na temeljne pravice.

Julija 2020 je na primer skupina na visoki ravni za umetno inteligenco[26] po šestmesečnem pilotnem izvajanju, v katero je bilo vključenih več kot 350 deležnikov, izdala Ocenjevalni seznam za zaupanja vredno umetno inteligenco (ALTAI). ALTAI pomaga organizacijam pri prostovoljnem samoocenjevanju, ali je sistem umetne inteligence zanesljiv in zaupanja vreden, ter zmanjšuje morebitna tveganja za uporabnike. Podpira podjetja in javne uprave pri zastavljanju pravih vprašanj, povezanih s sedmimi zahtevami za odgovorno umetno inteligenco, ki so opredeljene v Etičnih smernicah za zaupanja vredno umetno inteligenco[27]. ALTAI se izrecno sklicuje na potrebo po izvedbi ocene učinka v zvezi s temeljnimi pravicami. Vključuje primere vprašanj za oceno učinka na enakost, pravico do zasebnosti, otrokove pravice, svobodo izražanja ter svobodo obveščanja in združevanja[28].

Več spletnih ocenjevalnih orodij je namenjenih uporabi umetne inteligence s strani javnih organov. Kanadska vlada je razvila orodje za oceno učinka algoritmov (angl. *algorithmic impact assessment tool*, AIA)[29], in sicer v skladu s kanadsko direktivo o avtomatiziranem odločanju[30]. AIA je postopek avtomatiziranega ocenjevanja, sestavljen iz več kot 50 vprašanj, ki se nanašajo na zahteve iz omenjene direktive. Vprašanja se nanašajo na področje temeljnih pravic, obravnavajo na primer vpliv sistema umetne inteligence na svobodo gibanja, verjetnost odvzema prostosti posamezniku, pravni status, dostop do financiranja ali drugih prejemkov in na avtohtone prebivalce. Vsakemu odgovoru se pripiše ocena, na podlagi katere se določi končna ocena, ki se nato objavi na spletni strani vlade.

Drug primer je zbirka orodij za presojo etičnih vidikov, Ethics Toolkit[31], ki je prosto dostopna in namenjena lokalnim oblastem. Temelji na pristopu obvladovanja tveganj, podpira pravične avtomatizirane odločitve in zmanjšuje nenamerno škodo posameznikom na področju kazenskega pravosodja, visokega šolstva, družbenih medijev in drugih področij.

Med nacionalnimi organi za človekove pravice velja omeniti danski Inštitut za človekove pravice, ki predlaga „hitro preverjanje skladnosti s človekovimi pravicami"[32]. To vključuje interaktivni spletni računalniški program, ki podjetjem omogoča, da izberejo ali spremenijo informacije v podatkovni bazi, ki ustrezajo njihovemu poslovanju ali področju delovanja, in preverijo spoštovanje pravic. Hitro preverjanje temelji na orodju za oceno skladnosti s človekovimi pravicami[33], ki deluje s pomočjo zbirke podatkov z več kot 350 vprašanji in 1000 ustreznimi kazalniki človekovih pravic. Kot referenčne vrednosti uporablja mednarodne standarde prava človekovih pravic. Ti se uporabljajo na vseh področjih delovanja in služijo kot smernice pri izvedbi ocene učinka tehnologije umetne inteligence.

Akademska dela so predlagala tudi operativne okvire za ocenjevanje tveganja pri uporabi tehnologije umetne inteligence. Nekatera se osredotočajo zlasti na ugotavljanje in obravnavanje posledic za temeljne pravice v zasebnem sektorju[34]. Druga se osredotočajo na razvoj etičnih in vrednotno naravnanih modelov (analiza družbenega vpliva uporabljenih podatkov) z ustanovitvijo *ad hoc* strokovnega (revizijskega) odbora[35].

Spet druga so razvila smernice za posebne študije primerov. Na primer, na področju kazenskega pravosodja okvir ALGO CARE[36] vzpostavlja postopek ocenjevanja korak za korakom za ovrednotenje ključnih pravnih in praktičnih pomislekov, ki bi jih bilo treba upoštevati v primerih, ko policija pri svojem delu uporablja algoritemska orodja za oceno tveganja.

Nekateri so se zavzemali za participativne načine vključevanja in upoštevanja stališč prizadetih imetnikov pravic in drugih skupnosti deležnikov pri razvoju ocene učinka ter javnega sodelovanja z njimi od samega začetka[37]. Drugi so pri oblikovanju praktičnih okvirov združili interdisciplinarno strokovno znanje s področij znanosti in prava[38].

5.2 IZVAJANJE OCENE UČINKA IN TESTIRANJE V PRAKSI

Praktično vsi sistemi, obravnavani v intervjujih, so bili predmet neke vrste testiranja, ki je vključevalo elemente ocene učinka. Vendar je šlo v večini primerov za tehnične ocene in ocene (učinka) v zvezi z varstvom podatkov. Le redko so bili obravnavani možni vplivi na druge temeljne pravice.

Nekateri anketiranci nasprotujejo izvedbi ocene učinka v zvezi s temeljnimi pravicami, ker po njihovem mnenju sistem nima negativnih učinkov na temeljne pravice oziroma o takšnih negativnih učinkih niso prepričani. Anketiranec s področja upravljanja prometa, ki uporablja kamere za spremljanje prometa, je denimo navedel, da so se testiranja nanašala samo na točnost sistema, ne pa tudi na temeljne pravice, razen spoštovanja predpisov o varstvu podatkov.

Nekateri anketiranci preprosto niso vedeli, ali so bile temeljne pravice upoštevane v okviru splošne ocene učinka, ki je bila izvedena.

Testiranje in razvojne faze

Precejšen del testiranja se opravi pred začetkom uporabe katerega koli novega sistema umetne inteligence. Kot so poudarili anketiranci, je prehod sistema umetne inteligence v produkcijsko okolje zelo zahtevna naloga. Kot je bilo že omenjeno, so predstavniki javne uprave in zasebna podjetja pri uporabi umetne inteligence praviloma precej previdni. Številni projekti, o katerih so anketiranci govorili, so še vedno v fazi razvoja ali v pilotni fazi, pri nekaterih niso še niti začeli s konkretnimi testiranji.

„Pri testiranju sistema v resnici nismo preverjali pravnih vidikov, preučevali smo le, ali je sistem dobičkonosen.“
(zasebno podjetje, Estonija)

Testiranje je mogoče opraviti v več fazah. Te vključujejo razvojno fazo (tako imenovana potrditev koncepta), pilotne faze pred uvajanjem, testiranja med uvajanjem in tudi po njem. Če je mogoče, se eksperimentiranje v živo izvaja v začetnih fazah, kar pogosto vključuje postopno uvajanje.

Ena od organizacij, ki je sodelovala v razgovorih in se ukvarja s testiranjem različnih aplikacij za podporo iskalcem zaposlitve, izvaja nenehna postopna testiranja. Izbrani člani organizacije testirajo orodje v resničnih situacijah, pri čemer uporabljajo kontrolne sezname. Anketiranec je povedal, da je prehod v fazo uvajanja poln izzivov in je načrtovano, da bo orodje predmet nadzora v realnem času.

V drugem primeru, ki vključuje avtomatizirano dodeljevanje socialnih prejemkov na podlagi vnaprej določenih pravil, so bile izvedene različne ocene. Pred začetkom uvajanja je skupina pravnikov, strokovnjakov za varstvo podatkov in socialne prejemke ter računovodij izvedla splošno oceno učinka. Potem je oddelek, odgovoren za uporabo sistema, opravil testiranje, na podlagi katerega se je moral odločiti, ali je sistem primeren za uporabo.

V nadaljevanju je potekalo spremljanje sistema pri njegovem uvajanju z uporabo pristopa po korakih. V prvem koraku je sistem sprejel približno polovico odločitev. V naslednjem koraku so bile sprejete odločitve samodejno razširjene na vse negativne odločitve. Potem je bilo dodano še eno področje odločb, vključno z vsemi odločbami o ukinitvi nadomestil. V času, ko so bili opravljeni razgovori, je bilo približno 95 % odločitev avtomatiziranih. Anketiranec je navedel, da so se med izvajanjem testov prepričali, da je sistem varen in da ni nikakršnih izrednih tveganj.

Podjetje, ki je razvijalo sistem za odkrivanje goljufij, je svoj sistem, ki temelji na pravilih, zamenjalo z orodjem za strojno učenje. Pred prehodom z enega sistema na drugega sta vzporedno delovala stari in novi sistem, da se preveri, ali je sistem strojnega učenja boljši od sistema, ki temelji na pravilih. Anketiranec je omenil, da so „v ozadju natančna analiza in neposredne povratne informacije, ki so pokazale, kolikšen bi bil vpliv na izgube glede na to, na koliko dobrih strank so negativno vplivali". Anketiranec je dodal, da so sistem strojnega učenja v celoti uvedli šele v trenutku, ko so „bili prepričani, da je [sistem strojnega učenja] v vseh vidikih boljši [kot sistem statičnih pravil]".

V drugih primerih uporabe ni obstajal noben predhodni avtomatizirani sistem in teste so pregledovali samo ljudje. Avtomatizirana aplikacija za prepis je bila na primer preizkušena med zaslišanji na sodišču, ko je to dovolil sodnik. Postopek je vključeval redne povratne informacije o pravilnosti zapisa iz aplikacije s strani sodnikov.

Eden izmed anketirancev s področja kazenskega pregona, ki dela na orodju za odkrivanje nasilja v družini, ugotavlja, da so pri uporabi sistema težave z natančnostjo in točnostjo. Če policist ni dovolj usposobljen in nima znanja o sistemu, kazalniki, ki jih zahteva sistem, ne bodo mogli zbrati zahtevanih informacij, kar bi lahko vodilo do napačnega rezultata. Poudarja, da se zanesljivost sistema preverja vsako leto, da se zagotovijo kakovost obeh uporabljenih vprašalnikov, popolnost podatkov in usposabljanje policistov, ki uporabljajo sistem umetne inteligence. Znotraj tega postopka se prav tako določi, kako se uporabljajo zakoni in protokoli o varstvu osebnih podatkov. Obravnavani testi se v večinski meri osredotočajo na tehnične vidike in splošno delovanje.

Temeljne pravice in ocene učinka v zvezi z varstvom podatkov

Razen varstva podatkov, ki so ga omenili vsi anketiranci, praviloma niso bile upoštevane druge temeljne pravice. Anketiranci so o drugih možnih vplivih na temeljne pravice in o morebitni presoji teh vidikov začeli razmišljati šele, ko jih je izpraševalec k temu izrecno spodbudil.

Številni anketiranci se na splošno zavedajo vprašanj, povezanih z diskriminacijo, vendar so o tem pogosto začeli govoriti šele potem, ko jih je izpraševalec izrecno vprašal o diskriminaciji. Kljub temu niso posredovali nobenih informacij o kakršnih koli formalnih, poglobljenih testiranjih za presojo diskriminacije. Na splošno so anketiranci izključili možnost, da njihov sistem diskriminira na podlagi zaščitenih osebnih okoliščin. Eden od anketirancev je navedel, da opravlja testiranje sistema z vidika zakonodaje s področja varstva podatkov in konkretnih pravnih predpisov, vendar ne z vidika temeljnih pravic. In čeprav je anketiranec razmislil o morebitni diskriminaciji, je to možnost kmalu izključil. Navedel je le, da je treba to vprašanje imeti v mislih pri razvoju bodočih tehnologij.

Vendar pa obstajajo primeri, v katerih je bilo v fazi testiranja sistemov umetne inteligence na splošno upoštevano tudi vprašanje varstva pred diskriminacijo. En anketiranec, zaposlen v organu občinske uprave, je povedal, da pravičnosti modela ne more oceniti, saj zaradi predpisov o varstvu podatkov ne more dostopati do podatkov, ki bi jih potreboval za takšno oceno. Po mnenju anketiranca „obstaja velika napetost glede splošne uredbe o varstvu osebnih podatkov. Zelo si želimo, da bi nam šlo dobro, v resnici pa stvari morda le še poslabšamo, saj se interpretacija podatkov potem izkaže za nemogočo nalogo.“

„Da, ocenjujemo zakonitost varstva osebnih podatkov in skladnost s posebnimi pravnimi akti na tem področju.“

(javna uprava, Estonija)

Večina anketirancev je navedla, da je bila izvedena ocena učinka v zvezi z varstvom podatkov, kot to zahteva zakon, čeprav v različnih oblikah. Banka je opravila testiranje orodja za analizo govora med klici, da bi izvedela več o ponavljajočih se težavah, in izvedla oceno učinka v zvezi z varstvom podatkov (DPIA) posebej zaradi testiranja tega orodja. Rezultat je bil, da je sistem mogoče testirati, če se podatki uporabijo samo v fazi testiranja in se po določenem obdobju izbrišejo in če je dostop do podatkov s strani zaposlenih omejen na fazo testiranja in nadzorovan. Za uporabo orodja v praksi bi bila v tem primeru potrebna ponovna presoja učinka v zvezi z varstvom podatkov.

Včasih ni jasno, v kolikšni meri je uporaba umetne inteligence in sorodnih tehnologij, zlasti uporaba algoritmov, primerna za oceno učinka v zvezi z varstvom podatkov. Na primer na področju napovednega policijskega dela je bilo izvedenih nekaj ocen učinka v zvezi z varstvom podatkov, ki pa so se nanašale na osnovno arhitekturo sistema in ne na konkretno orodje umetne inteligence. Drug anketiranec, ki se ukvarja z uporabo algoritmov v finančnih storitvah, je omenil tudi, da orodja za strojno učenje kot takega ne ocenjujejo v okviru ocene učinka v zvezi z varstvom podatkov, ker se po njihovem mnenju ta postopek ne nanaša na sistem strojnega učenja (temveč na uporabljene osnovne podatke).

Eden izmed anketirancev je menil, da ocena učinka v zvezi z varstvom podatkov za toplotni zemljevid kaznivih dejanj ni dovolj poglobljena, da bi zaščitila kakovost modela, in da sistem ni opremljen za medsektorsko uporabo podatkov, kjer bi lahko veljala različna pravila. Navedel je, da so potrebni dodatni standardi.

Anketiranec, ki se ukvarja z upravljanjem migracij, je navedel, da so v analizo vključeni pooblaščenci za varstvo podatkov. Pravna služba ima specializirano orodje umetne inteligence za nadzor kakovosti, ki omogoča preučevanje vidikov varstva podatkov v njihovem sistemu. Vendar pa je anketiranec omenil tudi, da je potrebnih še več usmeritev.

Podjetja, aktivna na področju ciljno usmerjenega oglaševanja, so preučila vsa vprašanja v zvezi z varstvom podatkov, čeprav nekateri anketiranci niso bili prepričani, ali je bila izvedena ocena učinka. Podjetja so na primer ocenjevala, ali se pri ciljno usmerjenem komuniciranju obravnava le osebe, ki so podale privolitev. Pri ciljno usmerjenih oglasih se je preverjalo, ali so informacije za morebitno ponovno identifikacijo izbrisane, vključno z vprašanjem, ali so piškotki in sledilniki anonimizirani.

Kar zadeva splošno oceno učinka v zvezi z varstvom podatkov, nekateri anketiranci niso znali točno odgovoriti na vsa vprašanja, saj to ni bilo njihovo področje odgovornosti. Drugi so vedeli, da obstaja pozitivna ocena učinka, vendar niso bili seznanjeni s podrobnostmi. Zdi se, da je ocena zakonitosti včasih ločena od tehničnih vidikov, pri čemer tehnično osebje ni seznanjeno z ocenami zakonitosti. Anketiranec iz zasebnega podjetja, ki se ukvarja z oceno kreditnega tveganja, je povedal: „Včasih predlagam, kako bi lahko razvili nek sistem, a potem mi vodja službe za skladnost pove, da ni v skladu z zakonodajo.“

Revizije in sodelovanje z zunanjimi (nadzornimi) organi

Vsi organi javne uprave in zasebna podjetja, ki so sodelovali v raziskavi FRA, so pred uvajanjem kakršnih koli orodij umetne inteligence izvajali

določeno testiranje. To je bilo pogosto povezano z obstoječimi notranjimi in zunanjimi postopki nadzora. Uporaba umetne inteligence je pogosto predmet notranjih revizijskih postopkov v podjetjih in organih javne uprave, čeprav nujno ne gre za formalizirane revizijske postopke. Nekateri anketiranci so omenili, da si prizadevajo za formalizacijo obstoječih notranjih revizijskih postopkov za nadzor sistemov umetne inteligence.

Anketiranci iz javnega sektorja pravijo, da morajo biti pred uvajanjem kakršnega koli orodja umetne inteligence za podporo pri odločanju še posebej previdni. Predstavnik, ki se ukvarja z upravljanjem migracij v javni upravi, navaja, da „v zasebnem sektorju lahko [napačni rezultati] vodijo do poslovnih izgub, pri delu policije pa lahko vplivajo na življenja ljudi in njihove temeljne pravice".

Vendar pa predstavniki organov javne uprave in podjetij včasih ne vedo, kdo je odgovoren za preverjanje in nadzor nad uporabo umetne inteligence. Zdi se, da so organi javne uprave z vidika nadzora njihovih sistemov umetne inteligence pod strožjim nadzorom. Ta se pogosto izvaja v okviru rednih revizij, npr. v povezavi s proračunskim pregledom.

Nekateri anketiranci iz organizacij javnega in zasebnega sektorja so poročali, da so njihovi sistemi umetne inteligence trenutno v fazi pregledovanja v okviru obstoječih pregledov informacijske tehnologije (npr. rednih pregledov zbirk podatkov), kar velja v primerih, ko niso predvidena posebna preverjanja, ki bi se posebej osredotočala na uporabo umetne inteligence. Poleg tega so anketiranci poročali tudi o certifikacijskih shemah, tipičnih za posamezne sektorje, ki proučujejo tudi uporabo umetne inteligence – na primer na področju zdravstva ali finančnih storitev.

Več anketirancev je navedlo, da so bili v stiku z organi za varstvo podatkov. Nekatera podjetja in organi javne uprave so pred uporabo svojega sistema umetne inteligence zaprosili organe za varstvo podatkov za dovoljenje ali pa so bili vsaj na splošno v stiku z njimi. Podjetje, aktivno na področju ciljno usmerjenega oglaševanja, je na primer omenilo razpravo o svoji uporabi osebnih podatkov, ki je potekala z nacionalnim organom za varstvo podatkov.

Strokovnjaki, s katerimi je bil opravljen razgovor za namene tega poročila, so nadalje poudarili pomen organov za varstvo podatkov pri nadzoru sistemov umetne inteligence v zvezi z uporabo osebnih podatkov. Vendar pa so strokovnjaki odločno poudarili, da organi za varstvo podatkov nimajo zadostnih sredstev za uspešno opravljanje te naloge, in sicer iz dveh razlogov. Organi za varstvo podatkov pogosto nimajo ustreznega strokovnega znanja na področju umetne inteligence[39]. Poleg tega so njihovi proračuni preobremenjeni, delovna obremenitev pa velika.

Stališča strokovnjakov se z vidika potrebe po dodatnih nadzornih organih in morebitne ustanovitve posebne institucije za področje umetne inteligence razlikujejo. A vsi se strinjajo o tem, da morajo obstoječi organi v okviru svojih pooblastil obravnavati tudi teme, povezane z uporabo umetne inteligence.

Organi za enakost in druge institucije za varstvo človekovih pravic pa so tisti, ki po mnenju nekaterih anketirancev izvajajo nadzor nad morebitno diskriminacijo pri uporabi umetne inteligence. Sodelujoči so poudarili, da morajo te institucije pridobiti strokovno znanje na tem področju, kajti le tako bodo lahko bolje prispevale k nadzoru nad umetno inteligenco. A podobno kot pri organih za varstvo podatkov to organom za enakost pomeni velik izziv, še posebej ob upoštevanju pomanjkanja sredstev.

Več anketirancev je med organe, ki potencialno zagotavljajo ustrezen nadzor nad uporabo umetne inteligence, uvrstilo tudi organe za varstvo potrošnikov. Anketiranec, zaposlen v maloprodajnem podjetju, vidi rešitev v ustanovitvi svetovalne agencije, s katero bi se bilo mogoče posvetovati o uporabi umetne inteligence pri inovacijah, ne da bi to pomenilo takojšnjo uvedbo preiskave. Trenutno se podjetje o morebitnih bodočih tržnih kampanjah raje posvetuje z organi za varstvo potrošnikov kot z organi za varstvo podatkov. Razlog tiči v bojazni, da bi organi za varstvo podatkov lahko uvedli preiskavo prizadevanj podjetja.

„Ne samo, da smo proaktivni pri zmanjševanju tveganj, ampak smo deležni tudi dodatnih revizij. Prav tako včasih opažamo, da so nekatere zakonsko predpisane revizije precej površne. V našem primeru to ni dobro, ker imamo veliko podatkov o strankah.“

(zasebno podjetje, Estonija)

V kontekstu nadzora so tisti, ki razvijajo ali uporabljajo umetno inteligenco, pa tudi strokovnjaki, večkrat omenili izzive pri razumevanju vpliva uporabe umetne inteligence. Kljub potrebi po angažiranju obstoječih nadzornih organov ostajajo odgovornosti za nadzor uporabe umetne inteligence z vidika temeljnih pravic nejasne.

5.3 OCENA UČINKA V ZVEZI S TEMELJNIMI PRAVICAMI V PRAKSI

Številni ključni akterji na področju temeljnih pravic so pozvali k izvedbi ocene učinka v zvezi s temeljnimi pravicami še pred uporabo sistemov umetne inteligence. V tem oddelku so izpostavljeni nekateri elementi, ki bi jih bilo mogoče vključiti v takšno oceno.

Ocena učinka v zvezi s temeljnimi pravicami je potrebna, ker se zahteva kontekstualna ocena. Razlog je v tem, da se posamezni načini uporabe umetne inteligence izrazito razlikujejo glede na zahtevnost, stopnjo avtomatizacije, morebitne napake in škodo ter obseg in področje uporabe. Bolj ko je sistem umetne inteligence zapleten, težje je oceniti njegov potencialni učinek.

Čeprav se vključene temeljne pravice glede na področje uporabe lahko razlikujejo, je treba pri vsaki uporabi umetne inteligence upoštevati celoten spekter pravic. Vendar je verjetno, da bo uporaba umetne inteligence zadevala nekatere pravice, na katere sistemi umetne inteligence vplivajo najpogosteje. Iz razprave v prejšnjem poglavju jasno izhaja, da pridejo pri vsaki uporabi umetne inteligence v poštev vprašanja, povezana z varstvom podatkov, varstvom pred diskriminacijo ter dostopom do učinkovitih pravnih sredstev in poštenim sojenjem.

Zato bi lahko bile naslednje horizontalne točke osnovno izhodišče pri preverjanju učinkov umetne inteligence na izbrane pravice.

— **Zakonita obdelava podatkov** mora biti potrjeno skladna s predpisi o varstvu podatkov.

Če se uporabljajo osebni podatki, se uporablja tudi celoten pravni okvir varstva podatkov. Na ta način se zagotovi, da je obdelava zakonita in ne krši pravic osebe do spoštovanja zasebnega in družinskega življenja ter varstva podatkov.

— Obdelava **ne bi smela povzročati neenakega obravnavanja ali voditi v diskriminacijo zaščitenih skupin**.

Ocenjevanje nediskriminatornosti mora biti v samem jedru ocenjevanja umetne inteligence. Tudi minimalne razlike lahko privedejo do tveganja za kršitve načela prepovedi diskriminacije ali takšno tveganje povečajo. Prikrajšanje je odvisno od narave (vrsta škode), resnosti (velikost škode) in pomena (koliko ljudi je postavljenih v slabši položaj v primerjavi z drugo skupino ljudi). Statistične ocene razlik med skupinami so pomembno orodje za oceno nepoštene in diskriminatorne uporabe umetne inteligence[40].

— Osebe, ki so podvržene uporabi umetne inteligence in sorodnih tehnologij, **bi morale imeti možnost pritožiti se in uporabiti učinkovita pravna sredstva**.

Obstajati bi morali dostopni načini, na katere bi se ljudje lahko pritožili zoper morebitne odločitve in učinkovito dostopali do pravnih sredstev. To zajema razpoložljivost informacij za obrazložitev odločitev.

Poleg tega se uporabljajo tudi druge ustrezne pravice iz Listine. Organi javne uprave, ki uporabljajo umetno inteligenco, morajo upoštevati načela dobrega upravljanja. Podjetja morajo upoštevati pravila s področja varstva potrošnikov.

Ob upoštevanju konkretnega področja uporabe so lahko pomembne tudi druge pravice. Med njimi so:

- pravica do socialne zaščite na področju socialnih prejemkov,
- pravica do svobode izražanja in obveščanja pri uporabi umetne inteligence za podporo moderiranju spletnih vsebin,
- pravica do zbiranja in združevanja pri uporabi tehnologije za prepoznavanje obraza v javnih prostorih,
- pravica do izobraževanja pri uporabi umetne inteligence v izobraževalnem sektorju,
- pravica do azila pri uporabi umetne inteligence za podporo pri upravljanju migracij,
- pravica do kolektivnih pogajanj in ukrepov pri uporabi umetne inteligence v gospodarstvu priložnostnih del,
- pravica do pravičnih in poštenih delovnih pogojev pri uporabi umetne inteligence na delovnem mestu,
- pravica dostopa do preventivnega zdravstvenega varstva pri uporabi umetne inteligence v zdravstvenih storitvah,
- pravica do domneve nedolžnosti in pravica do obrambe pri uporabi umetne inteligence v pravosodnem sektorju ali za namene kazenskega pregona.

Informacije, potrebne za oceno morebitnega učinka v zvezi s temeljnimi pravicami pred uvajanjem umetne inteligence

Glede na raznolikost orodij, namenov in področij uporabe so ocene kontekstualne. Da bi se lahko smiselno odzvali na zgoraj navedene horizontalne točke in podali oceno glede na posebne pravice, povezane z različnimi primeri uporabe, morajo biti na voljo vsaj naslednje informacije:

- opis namena in konteksta uporabe sistema ter pravna podlaga;
- opis morebitne škode pri uporabi sistema, vključno z vprašanji v zvezi z lažno pozitivnimi ali lažno negativnimi rezultati in drugimi možnimi oblikami škode zaradi avtomatizacije in obsega uporabe;
- opis uporabljene tehnologije. Vključuje informacije o podatkih, uporabljenih pri izgradnji sistema, in pravni podlagi za obdelavo teh podatkov. Opis ustreznih informacij, ki morajo biti vključene, je naveden v dokumentu FRA o kakovosti podatkov in umetni inteligenci[41];
- z dokazi podkrepljen opis pravilnosti delovanja sistema umetne inteligence v smislu rezultatov, ki temeljijo na učnih podatkih ter morebitnem testiranju in eksperimentih v dejanskih razmerah, če je primerno. V tem primeru je treba lažno pozitivne in lažno negativne rezultate obravnavati ločeno. Ti bi morali vključevati razčlenitve za čim več skupin, da se omogoči preverjanje morebitne diskriminacije (npr. razlike v pravilnosti med ženskami in moškimi);
- če so že na voljo, tudi informacije o skladnosti z obstoječimi standardi in morebitnih pridobljenih certifikatih.

Naknadne ocene in zaščitni ukrepi

Nazadnje bi k uporabi umetne inteligence na način, ki je skladen s temeljnimi pravicami, dodatno prispevali tudi naknadni zaščitni ukrepi. Ti lahko vključujejo:

- redno ponavljanje ocen po zaključenem uvajanju, kjer je to primerno. To je pomembno za pridobivanje informacij o morebitnih povratnih zankah in v primeru posodobitve pravil. V tem smislu je potrebno tudi beleženje informacij o uporabi in rezultatih sistema v obsegu, ki je dopusten z vidika varstva podatkov;
- poskrbeti, da se ljudje, na katere se nanaša uporaba umetne inteligence, tega zavedajo, saj drugače ne morejo izpodbijati nobene odločitve, ki vpliva na njih;
- razpoložljivost enostavno dostopnih kanalov za učinkovite pritožbe zoper odločitve, sprejete v okviru sistemov umetne inteligence.

Vključevanje zunanjih strokovnjakov, deležnikov in nadzornih organov

Zgornje informacije bi lahko predstavljale podlago za izvedbo posvetovanja z različnimi deležniki in strokovnjaki, preden se začne sistem umetne inteligence uporabljati. Glede na naravo konkretne aplikacije in njeno pravno podlago bi posvetovanje z ustreznimi deležniki zagotovilo, da ne bo spregledana nobena morebitna škoda in da se v oceno vključijo različni vidiki. Deležniki bi lahko vključevali civilno družbo, različne javne in zasebne organizacije ter strokovnjake z različnih področij temeljnih pravic, vključno z varstvom podatkov.

Kot je poudarilo deset strokovnjakov, s katerimi je bil opravljen razgovor v okviru tega poročila, so za nadzor umetne inteligence v okviru svojih pooblastil odgovorni tudi že obstoječi nadzorni organi. Razgovori kažejo, da so tudi sektorski organi in certifikacijske sheme do neke mere aktivni v tem smislu, na primer na področju zdravstvenega varstva in finančnega nadzora.

Pri spremljanju, razumevanju in učinkovitem odzivu na morebitne učinke umetne inteligence na širok spekter temeljnih pravic bi lahko imeli pomembno vlogo organi za varstvo podatkov, organi za enakost, institucije varuha človekovih pravic in nacionalne institucije za človekove pravice, ki bi prispevali in nadzorovali dano področje ob upoštevanju različnih strokovnih vidikov. Vendar sta, kot so anketiranci navajali v razgovorih, v tem smislu potrebna obsežno izpopolnjevanje in zagotavljanje sredstev.

Končne opombe

1 Svet Evrope, Komisar za človekove pravice (2019), *Unboxing Artificial Intelligence: 10 steps to protect Human Rights – Recommendation* (Odkrivanje umetne inteligence: 10 korakov za varstvo človekovih pravic – priporočilo), Svet Evrope, Strasbourg, maj 2019.

2 Janssen, H. L. (2020), „An approach for a fundamental rights impact assessment to automated decision-making" (Pristop k oceni tveganja človekovih pravic pri avtomatiziranem odločanju), *International Data Privacy Law*, zvezek 10, izdaja 1, februar 2020, str. 76–106; Mantelero, A., „AI and Big Data: A blueprint for a human rights, social and ethical impact assessment" (Umetna inteligenca in masovni podatki: načrt za oceno učinkov na človekove pravice, socialo in etiko), *Computer Law & Security Review*, zvezek 34, izdaja 4, avgust 2018, str. 754–772; Edwards, L. in Veale, M. (2017), „**Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For**" (Sužnji algoritmov? Zakaj pravica do obrazložitve najverjetneje ni sredstvo, ki ga iščete), 23. maj 2017, 16 *Duke Law & Technology Review* 18.

3 Access Now (2020), ***Access Now's submission to the Consultation on the "White Paper on Artificial Intelligence - a European approach to excellence and trust"*** (Prispevek organizacije Access Now k razpravi o Beli knjigi o umetni inteligenci – evropski pristop k odličnosti in zaupanju).

4 Svet Evrope, ***Priporočilo Sveta Evrope CM/Rec (2020)1 Odbora ministrov državam članicam o posledicah algoritemskih sistemov za človekove pravice***, 8. april 2020, odst. 5.2. (ocena učinkov na človekove pravice).

5 Za podrobno razpravo o varstvu pred diskriminacijo glej: Wachter S., Mittelstatt, B. in Russel C. (2020), ***Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI*** (Zakaj pravičnosti ni mogoče avtomatizirati: premostitev vrzeli med protidiskriminacijsko zakonodajo EU in umetno inteligenco).

6 Vlada Kanade (2019), **Direktiva o avtomatiziranem odločanju**.

7 Združeni narodi, ***Vodilna načela Združenih narodov o podjetništvu in človekovih pravicah***, potrjena z Resolucijo Sveta za človekove pravice 17/4, A/HRC/RES/17/4 z dne 6. julija 2011, načela 18, 19 in 20.

8 Janssen, H. L., „**An approach for a fundamental rights impact assessment to automated decision-making**" (Pristop k oceni tveganja človekovih pravic pri avtomatiziranem odločanju), *International Data Privacy Law*, zvezek 10, izdaja 1, februar 2020, str. 76–106.

9 Strokovna skupina na visoki ravni za umetno inteligenco (2019), ***Etične smernice za zaupanja vredno umetno inteligenco***, 8. april 2019, poglavje III.

10 Prav tam, str. 15.

11 Glej npr: IBM, ***Everyday Ethics for Artificial Intelligence*** (Vsakodnevna etika pri uporabi umetne inteligence), 2019; Sony*,* ***Sony Group AI Ethics Guidelines*** (Etične smernice za uporabo umetne inteligence skupine Sony), 2019; Vodaphone, ***Vodaphone's AI framework*** (Okvir za uporabo umetne inteligence družbe Vodaphone), 2019; Arborus International in Orange, ***International Charter for Inclusive AI*** (Mednarodna listina za vključujočo umetno inteligenco), 21. april 2020, ki jo je podpisalo več kot 40 zasebnih podjetij, vključno s podjetji Camfil, Danone, EDF, L'Oréal, Metro, Sodexo itd.

12 Združenje Information Technology Industry Council (ITI), ***ITI AI Policy Principles*** (Načela politike za umetno inteligenco ITI), 2017.

13 ECP Platform for the Information Society, ***Artificial Intelligence Impact Assessment*** (Ocena učinkov umetne inteligence), Nizozemska, 14. november 2019.

14 Amnesty International, Access Now, Human Rights Watch, Wikimedia Foundation*,* ***The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems*** (Torontska deklaracija: varstvo pravic do enakosti in nediskriminacije pri uporabi sistemov strojnega učenja), 16. maj 2018 na **RightsCon Toronto**; Univerza v Montrealu, *Montreal Declaration Responsible AI* (Montrealska deklaracija o odgovorni umetni inteligenci), 2018.

15 Inštitut inženirjev elektrotehnike in elektronike (IEEE), Globalna pobuda za etiko avtonomnih in inteligentnih sistemov, ***Ethically Aligned Design: Prioritizing Human Wellbeing with Autonomous and Intelligent Systems*** (Etično usklajena zasnova: prednostna obravnava blaginje ljudi pri uporabi avtonomnih in inteligentnih sistemov)*,* 2019; Future of Life Institute, ***Asilomar AI Principles*** (Načela umetne inteligence iz Asilomarja), rezultati druge konference inštituta Future of Life Institute o prihodnosti umetne inteligence, 2017.

16 Glej na primer: ECP Platform for the Information Society, ***Artificial Intelligence Impact Assessment*** (Ocena učinkov umetne inteligence)*,* Nizozemska, 14. november 2019; pobuda IEEE.

17 Association for Computer Machinery (ACM), ***ACM Code of Ethics and Professional Conduct*** (Kodeks etike in poklicnega ravnanja ACM), 22. junij 2018.

18 Future of Humanity Institute, Univerza v Oxfordu, ***Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development*** (Standardi za upravljanje umetne inteligence: mednarodni standardi za globalno usklajevanje na področju raziskav in razvoja umetne inteligence), april 2019.

19 ISO, *standardi ISO/IEC JTC 1/SC 42, Umetna inteligenca, Standard in/ali projekt pod neposredno odgovornostjo sekretariata ISO/IEC JTC 1/SC 42*, ISO, ***ISO/IEC TR 24028:2020 Informacijska tehnologija – Umetna inteligenca – Pregled zanesljivosti umetne inteligence****,* maj 2020. Med drugim določa „pristope k ocenjevanju in doseganju razpoložljivosti, odpornosti, zanesljivosti, natančnosti, varnosti, zaščite in zasebnosti sistemov umetne inteligence". Drugi standardi ISO, ki so septembra 2020 še v razvoju: ISO/IEC CD 23894 Informacijska tehnologija – Umetna inteligenca – Smernice za obvladovanje tveganj, ISO/IEC AWI TR 24027 Informacijska tehnologija – Umetna inteligenca (UI) – Pristranskost v sistemih UI in odločanje, podprto z UI, ter ISO/IEC AWI TR 24368 Informacijska tehnologija – Umetna inteligenca – Pregled etičnih in družbenih vprašanj, več informacij na voljo na **spletnem mestu ISO**; Inštitut inženirjev elektrotehnike in elektronike (IEEE), IEEE *P7003™ Pristranskost algoritmov*; Nemško zvezno združenje za umetno inteligenco (*KI Bundesverband*), *KI Bundesverband Guetesiegel* (Pečat kakovosti zveznega združenja za umetno inteligenco), 22. marec 2019.

20 Delovna skupina iz člena 29, ***Smernice glede ocene učinka v zvezi z varstvom podatkov***, wp248rev.01, 13. oktober 2017, Priloga 2 **–** Merila za sprejemljivost ocene učinka v zvezi z varstvom podatkov.

21 Strokovna skupina na visoki ravni za umetno inteligenco, ***Ocenjevalni seznam za zaupanja vredno umetno inteligenco (ALTAI) za samoocenjevanje***, 17. julij 2020.

22 Vlada Kanade, ***Algorithmic Impact Assessment Tool*** (Orodje za ocenjevanje učinka algoritmov), 2019; Danski inštitut za človekove pravice, ***Human rights compliance assessment quick check*** (Hitro preverjanje ocene skladnosti s človekovimi pravicami), 7. junij 2016.

23 Center of Government Excellence, Univerza Johns Hopkins, ***Ethics & Algorithm toolkit*** (Zbirka orodij za etiko in algoritme), 2018; Vlada Kanade, ***Algorithmic Impact Assessment Tool*** (Orodje za ocenjevanje učinka algoritmov), 2019.

24 Delovna skupina iz člena 29, ***Smernice glede ocene učinka v zvezi z varstvom podatkov****,* wp248rev.01, 13. oktober 2017, Priloga 2 (poudarek na varstvu podatkov); Danski inštitut za človekove pravice, ***Human Rights Impact Assessment Guidance and Toolbox*** (Smernice in zbirka orodij za oceno učinkov na človekove pravice), 2016; AI Pulse, ***Creating a Tool to Reproducibly Estimate the Ethical Impact of Artificial Intelligence*** (Ustvarjanje orodja za ponovljivo oceno etičnega učinka umetne inteligence), 26. september 2019.

25 Fairness, Accountability, and Transparency in Machine Learning (FAT/ML), ***Principles for Accountable Algorithms and a Social Impact Statement for Algorithms*** (Načela za zasnovo odgovornih algoritmov in izjava o družbenem vplivu algoritmov), 2019.

26 Strokovna skupina na visoki ravni za umetno inteligenco, ***Ocenjevalni seznam za zaupanja vredno umetno inteligenco (ALTAI) za samoocenjevanje***, 17. julij 2020.
27 Glej npr. človeško posredovanje in nadzor; tehnično zanesljivost in varnost; zasebnost in upravljanje podatkov; preglednost; raznolikost, varstvo pred diskriminacijo in pravičnost; družbeno in okoljsko blaginjo; odgovornost. Strokovna skupina na visoki ravni za umetno inteligenco (2019), ***Etične smernice za zaupanja vredno umetno inteligenco***, 8. april 2019.
28 Strokovna skupina na visoki ravni za umetno inteligenco, ***Ocenjevalni seznam za zaupanja vredno umetno inteligenco (ALTAI) za samoocenjevanje***, 17. julij 2020, str. 5.
29 Vlada Kanade, ***Algorithmic Impact Assessment tool*** (Orodje za ocenjevanje učinka algoritmov), 2019.
30 Vlada Kanade, **Direktiva o avtomatiziranem odločanju**, 2019, člen 6 in Dodatek C.
31 Center of Government Excellence, Univerza Johns Hopkins, ***Ethics & Algorithm toolkit*** (Zbirka orodij za etiko in algoritme), 2018.
32 Danski inštitut za človekove pravice, ***Human rights compliance assessment quick check*** (Hitro preverjanje ocene skladnosti s človekovimi pravicami), 7. junij 2016.
33 Danski inštitut za človekove pravice, ***Human Rights Impact Assessment Guidance and Toolbox*** (Smernice in zbirka orodij za oceno učinkov na človekove pravice), 2016.
34 Janssen, H. L., „**An approach for a fundamental rights impact assessment to automated decision-making**" (Pristop k oceni tveganja človekovih pravic pri avtomatiziranem odločanju), *International Data Privacy Law*, zvezek 10, izdaja 1, februar 2020.
35 Mantelero, A., „**AI and Big Data: A blueprint for a human rights, social and ethical impact assessment**" (Umetna inteligenca in masovni podatki: načrt za oceno učinkov na človekove pravice, socialo in etiko), *Computer Law & Security Review*, zvezek 34, izdaja 4, avgust 2018, str. 754–772; AI Pulse – Program on Understanding Law, Science, and Evidence (PULSE), UCLA School of Law, ***Creating a Tool to Reproducibly Estimate the Ethical Impact of Artificial Intelligence*** (Ustvarjanje orodja za ponovljivo oceno etičnega učinka umetne inteligence), 26. september 2019. Ta model vključuje vrsto vprašanj za oceno učinka projektov, ki jih omogoča umetna inteligenca, na človekove pravice.
36 Oswald, M., Grace, J., Urwin, S. in Barnes, G. C., ***Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental' proportionality*** (Modeli nadzora nad oceno tveganja pri uporabi algoritmov: lekcije iz modela Durham HART in eksperimentalna proporcionalnost), *Information & Communications Technology Law*, zvezek 27, 2018 – **izdaja 2**.
37 AINOW, *Algorithmic Impact Assessments: a Practical Framework for Public Agency Accountability* (Algoritemske ocene učinka: praktični okvir za odgovornost javnih agencij), april 2018.
38 The Institute for Ethical AI & Machine Learning (Ethical ML Network (BETA)), ***The Machine Learning Maturity Model*** (Zrelostni model strojnega učenja), 2019.
39 Brave (2020), ***Europe's governments are failing the GDPR*** (Evropske vlade ne spoštujejo GDPR).
40 Glej Wachter S., Mittelstatt, B. in Russel C. (2020), ***Why fairness cannot be automated: bridging the gap between EU non-discrimination law and AI*** (Zakaj pravičnosti ni mogoče avtomatizirati: premostitev vrzeli med protidiskriminacijsko zakonodajo EU in umetno inteligenco).
41 FRA (2019), ***Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*** (Kakovost podatkov in umetna inteligenca – blažitev pristranskosti in napak pri varstvu temeljnih pravic), Luxembourg, Urad za publikacije, junij 2019.

6.

POGLED NAPREJ: IZZIVI IN PRILOŽNOSTI

To poročilo je objavljeno v času tekočega razvoja evropske zakonodaje in politike na področju umetne inteligence in globalnega boja proti koronavirusu. Pandemija covida-19 je potencialno pospešila sprejemanje inovativnih tehnologij. Vendar je tudi pokazala, da umetna inteligenca ni čudežna rešitev za vse težave in da njeno uvajanje prinaša najrazličnejše izzive.

To poročilo jasno kaže, da uporaba sistemov umetne inteligence posega na področje širokega spektra temeljnih pravic. Prav tako razkriva, da številna podjetja in organi javne uprave že uporabljajo ali načrtujejo uporabo umetne inteligence in sorodnih tehnologij. Vendar pri teh tehnologijah govorimo o različnih ravneh kompleksnosti. V večini primerov se uporabljajo relativno preprosti algoritmi. Tudi stopnja avtomatizacije je različna. V večini primerov je sprejemanje odločitev predmet človeškega nadzora, obstajajo pa tudi izjeme.

Trenutno uporabljene aplikacije so pogosto šele v fazi razvoja. Zakonodajalci EU in nacionalni zakonodajalci ter oblikovalci politik bi morali imeti to realnost v mislih – zlasti ko so soočeni z optimističnimi pričakovanji v zvezi s potencialom umetne inteligence v primerjavi z izzivi, ki so povezani z uporabo novih tehnologij, in potrebo po njihovi ureditvi.

„Poskušamo gledati v prihodnost. Vedno več bo avtomatizacije.“
(zasebno podjetje, Estonija)

Velika večina predstavnikov javne uprave in podjetij načrtuje, da bo še naprej delala na razvoju umetne inteligence ali jo uporabljala. Samo dva anketiranca sta navedla, da ne bosta več uporabljala ali razvijala umetne inteligence. Druga sogovornika pa sta previdna. Želita počakati in opazovati, kaj počnejo drugi, kot enega od razlogov pa sta navedla tudi pomanjkanje sredstev za kritje dela, povezanega z uvajanjem umetne inteligence.

„Naslednji koraki so povezani s preglednostjo in odprtimi podatki: to pomeni, da se objavijo ne le informacije v formatu pdf, ampak tudi informacije v ponovno uporabljivi obliki, tako da se omogoči njihova ponovna uporaba tako interno kot tudi znotraj celotnega zasebnega sektorja.“
(javna uprava, Španija)

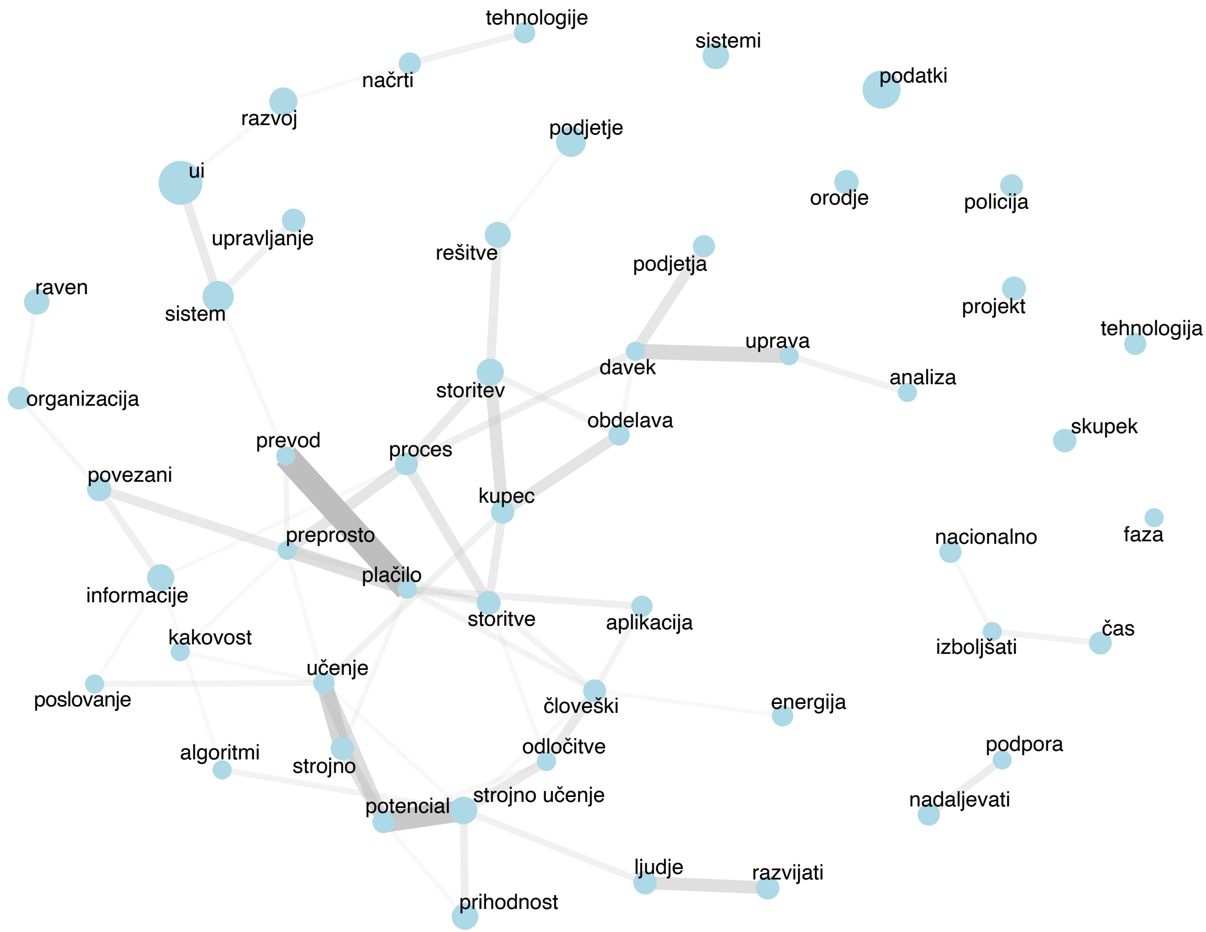
Večina pa je dejala, da bodo še naprej razvijali ali testirali orodja in (podatkovno) infrastrukturo v zvezi z uporabo umetne inteligence. To vključuje začetek novih ali nadaljevanje tekočih pilotnih projektov, ocenjevanje obstoječih prizadevanj, izmenjavo podatkov in rezultatov z drugimi, povečanje kakovosti podatkov ali poskuse pridobivanja drugih virov podatkov.

Nekateri anketiranci so omenili, da sodelujejo v tekočih razpravah, in izrazili željo prispevati k nadaljnjemu razvoju zakonodaje. Še vedno menijo, da predstavlja sedanji položaj – odsotnost usklajene zakonodaje na tem področju – oviro za nadaljnjo uporabo umetne inteligence. Poleg tega so nekateri anketiranci povedali, da se ukvarjajo z vprašanji, ki se nanašajo na razložljivost umetne inteligence. To pomeni, da delajo na metodah, ki podpirajo razumevanje in razlago odločitev, sprejetih na podlagi bolj zapletenih oblik umetne inteligence. Nekateri so pokazali željo, da se podrobneje seznanijo z etičnimi in pravnimi vidiki.

„Umetna inteligenca je odlična stvar, vendar se jo moramo naučiti uporabljati.“
(zasebno podjetje, Španija)

Slika 7 prikazuje korelacije besed, ki so jih anketiranci pogosto uporabljali, ko so govorili, kakšna bo njihova prihodnja uporaba umetne inteligence. Na sliki so prikazane teme, ki so se pogosto pojavljale. Anketiranci so na primer v razpravi o prihodnjem razvoju pogosto uporabljali besedo „podatki“.

SLIKA 7: KORELACIJE BESED, KI SO JIH ANKETIRANCI POGOSTO UPORABLJALI, KO SO GOVORILI O NAČRTIH ZA UPORABO UMETNE INTELIGENCE V PRIHODNOSTI



Opombe: Zasnovano na podlagi besedila iz povzetkov intervjujev, ko so anketiranci govorili o tem, kako nameravajo v prihodnosti uporabljati umetno inteligenco. Vključene so besede, omenjene vsaj desetkrat. Debelina črte, ki povezuje besede, označuje, kako močna je povezava med konkretnimi besedami v besedilu. Velikost pik označuje pogostost uporabljenih besed.

Vir: FRA, 2020

Učinkovito in ustrezno varstvo temeljnih pravic v EU je ključen cilj sedanjih prizadevanj za boljšo ureditev uporabe umetne inteligence. V okviru nastajanja prihodnje zakonodaje EU o umetni inteligenci bela knjiga Evropske komisije obravnava trenutne vrzeli in pomaga zmanjšati negotovost glede uporabe umetne inteligence v zvezi s temeljnimi pravicami ter povečati preglednost in odgovornost uporabe umetne inteligence v smislu temeljnih pravic. Vključuje zahteve za uporabo umetne inteligence, ki so neposredno povezane z informacijami, potrebnimi za oceno učinka umetne inteligence v zvezi s temeljnimi pravicami, kot je navedeno zgoraj.

Pri ocenjevanju in varstvu temeljnih pravic so izjemno pomembne zahteve, povezane z opisom učnih podatkov, podatki samimi in vodenjem evidenc, informacijami, ki jih je treba zagotoviti osebam, na katere vpliva umetna inteligenca, zanesljivostjo in natančnostjo ter človeškim nadzorom. V zvezi s tem dokazno gradivo, predstavljeno v tem poročilu, ponuja splošen vpogled v to, kako lahko različne tehnologije vplivajo na temeljne pravice in kakšni zaščitni ukrepi so potrebni za zagotovitev takšne uporabe umetne inteligence, ki bo v praksi v skladu s temeljnimi pravicami.

Hkrati bodo nadaljnje raziskave posledic uporabe umetne inteligence za temeljne pravice na posebnih področjih še naprej podpirale politična in zakonodajna prizadevanja na ravni EU, katerih cilj je širše oblikovanje digitalne prihodnosti Evrope. FRA bo še naprej preučevala temeljne posledice uporabe umetne inteligence prek izvajanja bolj osredotočene analize posameznih primerov uporabe. Z namenom širitve znanja o tem, kaj bi v konkretnih primerih lahko šlo narobe, in posledično podpore pri zmanjševanju in preprečevanju kršitev temeljnih pravic, bo FRA preučila tudi morebitne simulacijske študije. Te lahko pokažejo, kako pristranski algoritmi lahko negativno vplivajo na temeljne pravice.

Uporaba umetne inteligence pogosto vključuje avtomatizacijo nalog, ki so jih prej izvajali ljudje. Pri tem moramo priznati, da človeško vedenje včasih ni v skladu s temeljnimi pravicami, tako pri uporabi umetne inteligence kot brez nje. Na primer, tudi policija se lahko ukvarja z nezakonitim profiliranjem. Odločitve organov javne uprave ali podjetij so lahko včasih posledica negativnih stereotipov.

Pri trenutnem razvoju uporabe umetne inteligence se je treba zavedati, da obstaja možnost diskriminacije v zvezi s podatki, na katerih je zgrajen sistem umetne inteligence, in v zvezi z osnovno predpostavko, da so v razvoj in uvajanje sistema vključeni ljudje. Avtomatizacija nekaterih nalog, ne da bi v celoti razumeli, kaj se avtomatizira, bi lahko privedla do nezakonite obdelave podatkov, uporabe tehnologije, ki nepošteno obravnava ljudi, in onemogočila izpodbijanje nekaterih rezultatov, če omenimo le nekatere izzive.

Vendar pa se lahko večja razpoložljivost podatkov in tehnoloških orodij uporabi tudi za boljše razumevanje, kje in kako prihaja do neenakega obravnavanja. Trenutni tehnološki razvoj in večja razpoložljivost podatkov sta tudi edinstvena priložnost za boljše razumevanje družbenih struktur, kar lahko pripelje do učinkovitejšega spoštovanja temeljnih pravic. Priložnosti, ki jih prinaša umetna inteligenca, lahko prispevajo tudi k boljšemu razumevanju in posledično zmanjševanju kršitev temeljnih pravic.

Stik z EU

Osebno
Po vsej Evropski uniji je na stotine centrov Europe Direct. Naslov najbližjega lahko najdete na spletu (**european-union.europa.eu/contact-eu/meet-us_sl**).

Po telefonu ali pisno
Europe Direct je služba, ki odgovarja na vaša vprašanja o Evropski uniji. Nanjo se lahko obrnete:
- s klicem na brezplačno telefonsko številko: 00 800 6 7 8 9 10 11 (nekateri ponudniki lahko klic zaračunajo),
- s klicem na navadno telefonsko številko: +32 22999696,
- z uporabo obrazca: **european-union.europa.eu/contact-eu/write-us_sl**.

Iskanje informacij o EU

Na spletu
Informacije o Evropski uniji v vseh uradnih jezikih EU so na voljo na spletišču Europa (**european-union.europa.eu**).

Publikacije EU
Publikacije EU si lahko ogledate ali naročite na **op.europa.eu/sl/publications**. Za več izvodov brezplačnih publikacij se obrnite na Europe Direct ali najbližji dokumentacijski center (**european-union.europa.eu/contact-eu/meet-us_sl**).

Zakonodaja EU in drugi dokumenti
Do pravnih informacij EU, vključno z vso zakonodajo EU od leta 1952 v vseh uradnih jezikovnih različicah, lahko dostopate na spletišču EUR-Lex (**eur-lex.europa.eu**).

Odprti podatki EU
Na portalu **data.europa.eu** lahko dostopate do odprtih zbirk podatkov institucij, organov in agencij EU. Zbirke podatkov lahko brezplačno prenesete ter jih ponovno uporabite za komercialne in nekomercialne namene.

FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

SPODBUJANJE IN VARSTVO VAŠIH TEMELJNIH PRAVIC V EU

Listina EU o temeljnih pravicah

Dostop do pravnega varstva

Nediskriminacija

Informacijska družba

FRA – AGENCIJA EVROPSKE UNIJE ZA TEMELJNE PRAVICE
Schwarzenbergplatz 11 – 1040 Dunaj – Avstrija
Tel.+43 158030-0 – Faks +43 158030-699

fra.europa.eu

facebook.com/fundamentalrights
twitter.com/EURightsAgency
linkedin.com/company/eu-fundamental-rights-agency

Urad za publikacije Evropske unije

REPUBLIKA SLOVENIJA
ZAGOVORNIK NAČELA ENAKOSTI