ABA
AMERICANBARASSOCIATION
Rule of Law Initiative

МЕТАМОРФОЗИС
Фондација за интернет и општество

ИСТРАЖУВАЊЕ ЗА ЕФЕКТОТ НА НОВИТЕ ТЕХНОЛОГИИ, СО ОСОБЕН ФОКУС НА ВЕШТАЧКАТА ИНТЕЛИГЕНЦИЈА, ВРЗ ЧОВЕКОВИТЕ ПРАВА НА ИНТЕРНЕТ И РАЗВИВАЊЕ ЕТИЧКИ СТАНДАРДИ ЗА ЗАШТИТА НА ЧОВЕКОВИТЕ ПРАВА НА ИНТЕРНЕТ ПРИ АВТОМАТСКО ДОНЕСУВАЊЕ ОДЛУКИ

Скопје, 2024

Издавач: Фондација за интернет и општество Метаморфозис – Скопје

ул. Франклин Рузвелт број 19, Скопје 1000, Северна Македонија; www.metamorphosis.org.mk

За издавачот: Бардил Јашари, извршен директор, Фондација „Метаморфозис“

Уредник: Мила Јосифовска Даниловска, програмски менаџер, Фондација „Метаморфозис“

Автор: проф. д-р Игор Камбовски, м-р Елена Стојанова

Тираж: електронско издание

Скопје, 2024 г.

СОДРЖИНА

Листа на акроними

ВИ – вештачка интелигенција

ЕУ – Европска унија

GDPR – General Data Protection Regulation

LLM – Language Learning Models

ОЕЦД - Организацијата за економска соработка и развој

ОН – Обединети нации

АДА – Прв дигитален асистент

GSM – Global System for Mobile Commuication

GPS – Global Positioning System

Wi-Fi – Wireless Fidelity

ПВЗЛП – Проценка на влијанието на заштитата на личните податоци

УИНТ – Универзитет за информатички науки и технологии

ФЕИТ – Факултет за електротехника и информациски технологии

ФИНКИ – Факултет за информатички науки и компјутерско инженерство

UNESCO - United Nations Educational, Scientific and Cultural Organization

AI – Artificial inteligence

Извршно резиме

Вештачката интелигенција (ВИ) и нејзината интензивна употреба се релативно нов феномен. Меѓутоа, степенот до кој може да влијае на секојдневниот живот на луѓето, на редизајнирањето на деловните процеси и начинот на функционирање и организирање на секојдневните општествени, економски, социјални, трговски, политички, демократски и други видови процеси, и тоа глобално и за исклучително кратко време, е огромен и е многу непредвидлив. Од што се состои суштинската промена? Компјутерите, информатичките и комуникациските технологии и процесите на дигитализација примарно претставуваа средства и начини кои го олеснуваа собирањето, складирањето и обработката на податоци, со цел да се зголеми квалитетот и брзината на работата на луѓето. Сепак, со брзиот развој и проширување на примената на вештачката интелигенција - благодарение на нап редокот на длабоките невронски мрежи, зголемувањето на количеството на податоци погодни за машинско учење и зголемената достапност на микропроцесори погодни за обемни нумерички пресметки – се овозможи на компјутерите, роботите и други автоматизирани производи да извршуваат активности кои претходно биле карактеристични само за луѓето. Ова им овозможува на роботите да бидат користени за вредни и хумани цели (како на пример - опасни активности во рудници или во вселената) или на компјутерите да донесуваат заклучоци врз основа на обработка на огромни количини на податоци (на пример - анализирање на снимки и наоди за да се обезбедат попрецизни медицински дијагнози) и да се предвидат одредени настани (како што се економските модели и трендови или временската прогноза). Гледано од тие аспекти, корисноста на вештачката интелигенција е неспорна и неприкосновена. Сепак, потребно е да се земе предвид и проблемот со намалување на бројот на работни места (на пр. во банкарството, јавната администрација, индустријата). Од друга страна, пак, во областа на воената индустрија сè уште преовладува идејата и перцепцијата за корисноста на вештачката интелигенција во однос на свесноста за ризиците што веќе се случиле или потенцијално можат да се случат во иднина. Овие ризици се поврзуваат со појавата и развојот на таканаречената супер-вештачка интелигенција, која самостојно и сè понезависно се подобрува и интензивно се надградува, истата донесува одлуки на свој начин и постои реална можност, дури и опасност, да излезе од контрола и регулација.

Потенцијалните опасности често се илустрираат со следниот мисловен експеримент. Да речеме дека имаме хипотетички систем чија цел е да произведува што е можно повеќе спојници. Системот може да сфати дека луѓето може да претставуваат проблем при најоптималното произведување на спојници, бидејќи токму луѓето би можеле, на пример, да го исклучат системот. Системот согледува дека неговото исклучување е лоша вест за неговата цел – произведување на спојници. Бидејќи сè што му е кажано да прави е да смисли како оптимално да произведува спојници, може да одлучи да ги отстрани сите луѓе, за да не му „пречат“. Иако овој мисловен експеримент можеби малку ги преувеличува опасностите на ова конкретно сценарио, тој не е само сатира. Тој опипливо укажува на потребата на исклучително внимание при неконтролираното развивање и пуштање во јавност на алатките базирани на вештачка интелигенција[1].

Понатаму, можни се проблеми со системите за поддршка на одлучувањето. Машините и апликациите кои користат вештачка интелигенција ја рефлектираат пристрасноста на нивните креатори или социјалните предрасуди вградени во податоците што се користат за обука на машините и апликациите, бидејќи вештачката интелигенција учи од податоците. Таквите податоци содржат моментни и тековни предрасуди, а машините само ќе ги институционализираат таквите тенденции. Пристрасноста или нетранспарентноста се

[1] Вештачката интелигенција – застрашувачки корисна, но и опасна алатка во човечките раце. Новинска агенција Мета.мк., преглед на 20.12.2023 год.

влошуваат со брзото проширување на бројот на информациски системи управувани од податоци кои влегуваат во процесот на донесување одлуки. Исто така, постои можност и закана да се прекршат или ограничат основните права (на пр. правото на приватност и заштита на личните податоци). Имено, електронскиот провајдер на услуги во приватниот сектор собира податоци од интеракции или трансакции со корисници и подоцна ги користи за издвојување вредности, но корисникот може да се согласи или да приговара на обработката на податоците, додека во јавниот сектор и јавната управа употребата на информатичките технологии обично е пропишана и уредена со закон, така што, како и во поглед на употребата на вештачката интелигенција и технологии за обработка на податоци, поединецот не секогаш има право на приговор.

Личните податоци станаа незаменлив дел од развојот и употребата на вештачката интелигенција, каде што служат и се користат како основа за обука на системите за вештачка интелигенција во препознавање шеми, обрасци, подобрување на точноста и овозможување персонализација што се користи во различни индустрии. Дури и по развојот, функционирањето на системите за вештачка интелигенција зависи од податоците со кои „се храни“ вештачката интелигенција, без кои тие не би можеле да ги применат, ниту да ги дадат посакуваните резултати и да го усовршат своето учење. Општата регулатива за заштита на податоците (GDPR)[2] поставува низа стандарди за да се обезбеди високо ниво на заштита на личните податоци во ЕУ, додека многу национални закони се обврзани да ги следат овие стандарди. Од друга страна, вештачката интелигенција сè уште е регулирана преку меки законски инструменти, таканаречен “soft law”, како што се Етичките упатства за доверлива вештачка интелигенција, креирани на национално и супранационално ниво.

[2] Општа регулатива за заштита на податоците (GDPR)

Методологија на истражувањето

Главната цел на ова истражување е да се процени влијанието на новите технологии, со особен фокус на развојот и користењето на вештачката интелигенција, врз човековите права. Во поширока смисла, наодите од истражувањето треба да придонесат кон заштитата на човековите права во политиките поврзани со вештачката интелигенција преку натамошно застапување и развивање на капацитетите на целните групи но и преку реализирање на активности за подигнување на јавната свест. Дополнително, резултат на ова истражување е и развивање етички стандарди за заштита на човековите права при креирањето политики што го уредуваат автоматското донесување одлуки.

Истражувањето беше спроведено во период ноември – декември 2023 година.

За потребите на спроведување на ова истражување главно е користен методот на анализа на секундарни извори на податоци.

Со анализа на европската легислатива е даден преглед на правната рамка со која се регулира развојот и користењето на вештачката интелигенција, предностите но и ризиците кои произлегуваат по човековите права од несоодветна примена на важечките норми и утврдените принципи за етичко развивање и користење на вештачката интелигенција.

Со анализа на претходно спроведени истражувања, издадени насоки, водичи, упатства дадени од страна на меѓународни професионални тела и институции, даден е преглед на начините на кои се дефинираат дигиталните права, односно, правото на пристап до интернет и недискриминација, правото на слобода на изразување и информации, правото на слобода на собирање, здружување и учество, правото на приватност и заштита на личните податоци и заштитата на деца и млади луѓе.

Еден од најголемите предизвици е заштитата на приватноста и личните податоци при користењето на вештачката интелигенција. Преку анализа на европските регулативи за заштита на личните податоци но особено анализа на Законот за заштита на личните податоци на Република Северна Македонија и одделните правилници кои произлегуваат од овој закон, направена е анализа на начинот на почитување на начелата за заштита на личните податоци при развојот на вештачката интелигенција, начинот на заштита на правото да се биде исклучен од автоматското донесување на одлуки и заштитата на правата кои ги гарантира Законот за заштита на личните податоци.

Со цел добивање на приказ за тоа како сите актери вклучени во развојот и користењето на вештачката интелигенција работат кон воспоставување на баланс помеѓу брзиот развој на технологијата, носењето на нови и прилагодувањето на постоечките политики и обезбедувањето на заштита на дигиталните права, употребен е методот на интервју со засегнати страни.

Прашањата се наменети за владините и јавните институции, независните регулатрони тела и агенции, образовните институции и приватните компании кои се вклучени во различни фази на развојот на вештачката интелигенција.

Како резултат на употребените методи и сознанијата до кои е дојдено, дефинирани се и етичките стандарди за развој и користење на вештачката интелигенција.

1. Вовед

Развојот на вештачката интелигенција доведе до глобален трансформативен дигитален напредок кој сè повеќе влијае на нашиот секојдневен живот, работа, семејство, здравје, спорт, дружење, приватност и друго. Прогресот е препознатлив преку новите технологии за мапирање, лоцирање и препознавање, паметни телефони со гласовно управување, лични дигитални помошници, препознавање на ракопис за испорака на пошта, транспорт со автономни возила, финансиско тргување, паметна логистика, автоматизирано договарање, филтрирање спам пораки, превод на јазици и многу повеќе. Напредокот на вештачката интелигенција, исто така, обезбедува големи придобивки за нашата социјална благосостојба во области како што се прецизната хирургија и медицината воопшто, еколошка одржливост, образованието, администрацијата, производството, енергетиката, трговијата и друго.

Производите и услугите базирани на вештачка интелигенција денес се во широка употреба: автономни возила, различни видови роботи, системи за биометриска идентификација и категоризација на поединци, системи за управување со сообраќајот, системи за снабдување со вода, електрична енергија, гас, греење, електрична енергија, системи во образованието наменети за оценување, банкарски системи за оценување на кредитниот рејтинг на поединци, системи на вештачка интелигенција при вработување и управување со работници, системи наменети за судството и органите за кривично гонење, системи наменети за властите за контрола на патните исправи, визи, азил, мигранти, системи наменети за демократски процеси (електронско гласање и сл.) и многу други системи.

Вештачката интелигенција веќе има директно влијание врз економијата, политиката, образованието, културата, демократијата и човековите права. Можеме само да претпоставуваме какво влијание ќе има вештачката интелигенција врз нашите животи во иднина. Нејзиниот развој и влегување во нашето секојдневие денес отвора низа нови прашања: од прашањето за правниот субјективитет и одговорност на роботите со вештачка интелигенција, до прашањето за заканите за човековите права и демократијата од системите за вештачка интелигенција. Бројот на човекови права кои се загрозени поради развојот и примената на вештачката интелигенција се зголемува од ден на ден. Неопходно е правото да одговори на тој предизвик и да ги заштити основните човекови права и слободи. Донесувањето на нови правила и измените на постојните треба да создаде правен систем кој успешно ќе ја заштити највисоката вредност водена од идејата за правда, морал и етика. Тој правен систем на национално и меѓународно ниво мора да биде составен од правни норми кои се меѓусебно усогласени и претходно договорени. Законот е должен да одговори на развојот на новите технологии и да ја ограничи можноста за нивна злоупотреба, и да ги заштити човековите права и слободи. Поради енормно брзиот развој на новите технологии, неопходно е што поскоро да се создадат законски прописи во оваа област. Токму поради оваа причина, во изминатите десетина години, бројни меѓународни и национални тела се занимаваа со одредени прашања од правната регулација на вештачката интелигенција, а особено со прашањето за заштита на човековите права од можни закани од системите за вештачка интелигенција. Советот на Европа и Европската унија (ЕУ) во изминатите неколку години усвоија низа документи поврзани со одредени аспекти од законското регулирање на вештачката интелигенција, вклучително и аспекти на заштитата на човековите права.

Сите досегашни активности на националните и меѓународните експерти укажуваат дека правниот систем што ќе ја регулира вештачката интелигенција во иднина мора да биде дел од глобалниот правен механизам кој ги регулира дигиталните технологии воопшто и мора да вклучува кохерентен сет на обврзувачки и необврзувачки правила, кој ќе го регулира на фер, морален и етички начин секојдневното користење на вештачката интелигенција во различни

области од животот и работата на луѓето. Заемната врска меѓу правото и вештачката интелигенција не е еднонасочна улица, односно не само што правото влијае на вештачката интелигенција, туку вештачката интелигенција исто така влијае на правото на различни начини. Во многу аспекти, вештачката интелигенција може да влијае на поинаков и подобар начин на примената на правото. Некои од тие начини се автоматско преведување, предвидување ризик, управување со ресурси, пополнување формулари и слично. Исклучително актуелната, интересна и инспиративна тема за меѓусебната врска меѓу вештачката интелигенција и човековите права не само што има практично влијание врз нашето секојдневие, туку квалитетот на законската регулатива во оваа област ќе влијае на одредување на степенот на достоинство и вистинска слобода на поединецот во сајбер-просторот. Човековите права и слободи, како универзална вредност, мора да бидат заштитени од можни закани од производи и услуги засновани на алгоритми, т.е. вештачка интелигенција. Најдобар начин тоа да се случи е да се развие корпус на правни норми, кои ќе гарантираат ефективно остварување на правата и слободите за сите поединци без никаква дискриминација поради различности.

1.1. Историјат на развојот на вештачката интелигенција: Клучни датуми и имиња

Идејата за „машина што мисли“ датира од античка Грција. Но, од доаѓањето на електронското комуницирање, складирање податоци и пресметување, поважни настани и пресвртници во еволуцијата на вештачката интелигенција се следните:

1950: Алан Тјуринг го објавува трудот „*Computing Machinery and Intelligence*“[3]. Во трудот, Тјуринг - познат по декодирањето на нацистичкиот код ЕНИГМА за време на Втората светска војна - предлага да одговори на прашањето „дали машините можат да размислуваат?“ и го воведува Туринг тестот[4] за да утврди дали компјутерот може да ја демонстрира истата интелигенција (или резултатите од истата интелигенција) како и човекот. Оттогаш се дебатира за вредноста на Туринговиот тест.

1956: Џон Мекарти го измислил терминот „Вештачка интелигенција“ на првата конференција за вештачка интелигенција на колеџот Дартмут. Подоцна истата година, Ален Њуел, Џеј Си Шо и Херберт Сајмон го создале Logic Theorist[5], првата софтверска програма за вештачка интелигенција.

1967: Френк Розенблат го направил Mark 1 Perceptron, првиот компјутер базиран на невронска мрежа што „учел“ со обиди и грешки. Само една година подоцна, МарвинИнски и Сејмур Пеперт објавуваат книга со наслов „Perceptrons“[6], која станува и најзначајно дело за работата на невронските мрежи.

1980-ти: Невронските мрежи кои користат алгоритам за заднинско пропагирање за да се обучуваат станаа широко користени во апликациите за вештачка интелигенција.

1997: Deep Blue на IBM[7] го победил тогашниот светски шампион во шах Гари Каспаров, во два последователни шаховски меча.

---

[3] Компјутерски машини и интелигенција. А. М. Тјуринг., преглед на 18.12.2023 год.
[4] Туринг тест. Станфорд енциклопедија на филозофија., преглед на 18.12.2023 год.
[5] Логички теоретичар објасни - Сè што треба да знаете. History Computer., преглед на 14.12.2023 год.
[6] Перцепции – Вовед во компјутерска геометрија | MIT Press, М.Ински, С. А. Пејперт., преглед на 14.12.2023 год.
[7] Deep Blue. Chess., преглед на 18.12.2023.

2011: IBM Watson платформата ги победила шампионите Кен Џенингс и Бред Ратер во логичната игра *Jeopardy*![8]

2015: Кинескиот суперкомпјутер Minwa на Baidu[9] користи посебен вид длабока невронска мрежа наречена конволутивна невронска мрежа за да идентификува и категоризира слики со повисока стапка на точност од просечниот човек.

2016: Програмата AlphaGo на DeepMind, напојувана од длабока невронска мрежа, го победи Ли Содол, светскиот шампион во играта Go, во натпревар од пет сета[10]. Победата е значајна со оглед на огромниот број можни потези што се развивале во текот на играта (над 14,5 трилиони по само четири потези!). Подоцна, Google го купил DeepMind за пријавени 400 милиони американски долари.

2023: Зголемувањето на големите јазични модели или LLM[11], како што е ChatGPT, создава огромна промена во перформансите на вештачката интелигенција и нејзиниот потенцијал. Со овие нови генеративни практики за вештачка интелигенција, моделите за длабоко учење може да бидат претходно обучени користејќи огромно количество сурови, необработени и неозначени податоци.

1.2. Дефиниција: Што претставува Вештачката интелигенција (ВИ)

Терминот „вештачка интелигенција“ значи систем заснован на машина кој може, за даден сет на цели дефинирани од човекот, да прави предвидувања, препораки или одлуки кои влијаат на реалните или виртуелните средини“[12].

Вештачката интелигенција (ВИ) претставува способност на компјутер или компјутерски контролиран робот да извршува задачи вообичаено поврзани со интелигентни суштества. Терминот често се применува на проект за развој на системи обдарени со интелектуални процеси карактеристични за луѓето, како што е способноста за расудување, откривање значења, генерализирање или учење од минатото искуство. Од развојот на дигиталниот компјутер во 1940-тите, беше докажано дека компјутерите можат да се програмираат да извршуваат многу сложени задачи - како што се откривање докази за математички теореми или играње шах - со големо знаење и успех. Сепак, и покрај континуираниот напредок во брзината на компјутерската обработка и капацитетот на меморијата, сè уште нема програми што можат да одговараат на целосната човечка флексибилност на поширок домени или во задачи кои бараат големо секојдневно знаење. Од друга страна, некои програми ги достигнаа нивоата на перформанси на човечки експерти и професионалци во извршувањето на одредени специфични задачи, така што вештачката интелигенција во оваа ограничена смисла се наоѓа во различни апликации за подготовка на медицинска дијагноза, компјутерски пребарувачи, препознавање глас или ракопис, и чет-ботови[13]. Во својата наједноставна појавна форма, вештачката интелигенција е поле кое комбинира компјутерска наука и робусни збирки на податоци за да овозможи решавање на проблеми. Вештачката интелигенција исто така опфаќа под-полиња на машинско учење и длабоко учење, кои често се споменуваат во врска со вештачката интелигенција, и кои

[8] Ватсон, шампион на Jeopardy! IBM., преглед на 18.12.2023 год.
[9] Вештачката интелигенција на суперкомпјутерот Baidu Minwa ги надминува Google, Microsoft и луѓето во препознавањето на слики. Ентони Катбертсон., преглед на 18.12.2023 год.
[10] Вештачка интелигенција: AlphaGo на Гугл го победи Go мастерот Lee Se-dol. BBS News., преглед на 18.12.2023 год.
[11] Language Learning Models.
[12] Вештачка интелигенција (ВИ). Стејт департментот на САД., преглед на 24.11.2023 год.
[13] Вештачка интелигенција. Британика., преглед на 22.12.2023 год.

претставуваат алгоритми кои се обидуваат да создадат експертски системи за предвидувања или класификации врз основа на влезните податоци[14].

Од нејзината појава, колку и да денес ни изгледа застарена и рудиментирана, вештачката интелигенција помина низ многу циклуси на развој, но дури и за најголемите скептици, објавувањето на ChatGPT на OpenAI се чини дека претставува огромен исчекор напред и пресвртница во нејзиниот развој. Во меѓувреме, експертската и лаичката јавност создадоа мноштво анализи и теории во поглед на вештачката интелигенција, кои се движат од теолошко-оптимистички и надежни, па се до научно-фантастични и апокалиптични. Сепак, меѓу сите тие толкувања и предвидувања, интересна е тезата дека зголемената употреба на вештачката интелигенција ќе го направи човештвото похумано, покорисно и посвесно, а искористувањето на вештачката интелигенција за извршување на вообичаени или репетитивни задачи во секојдневието ќе му овозможи на човекот да се посвети на покреативни апликации и активности во работата и животот.

[14] Што е вештачка интелигенција (ВИ)? IBM., преглед на 20.11.2023 год.

2. Правно регулирање на вештачката интелигенција – Европска легислатива

Секоја нова освоена или барем допрена граница, во сите сфери на општеството, се карактеризира со недостаток на закони кои би ја уредиле таа „новоосвоена територија“. Во минатото творците на законите морале да се соочат и да се покорат на предизвиците кои произлегувале од конкретните општествени услови и конкретната заедница или држава. Меѓународното право се стреми кон премостување на празнините, но често биле потребни години и децении за да се оформи, договори и да стапи во сила конкретна правна норма. Во меѓувреме, интернетот се разви до невидени размери и длабоко навлезе во нашите животи, работа, во политиката и економијата, во воената индустрија и во комуникациите, со стремеж да ги редефинира основните сфаќања и природните права на човекот и човештвото, дефинирани низ историјата во мноштво повелби, декларации и меѓународни правни акти. Предизвикот да се регулира или, што е уште потешко, да се предвиди и да се регулира следниот технолошки исчекор, е голем и мултидисциплинарен. Создавањето на правото е комплицирано и бавен процес, а технолошките достигнувања се премногу брзи, така што се поставува прашање и дилема: дали правото може да ги следи динамичните промени и развој на информатичко-комуникациските технологии? Секој нов закон кој правно ги уредува односите кои настануваат со користење на новите технологии за брзо време застарува и се јавува потреба од негово ревидирање и изменување, и така во бесконечност. Ова придонесува за создавање на правни норми кои, за да можат да потраат, се прешироко поставени или непрецизни, што од своја страна ја зголемува правната несигурност што, пак, обратно пропорционално влијае на довербата кај субјектите кои учествуваат во е-трговијата, социјалните мрежи и воопшто, во сите комуникации и трансакции овозможени и спроведени преку. Исто така, техничките достигнувања создаваат одредени практични стандарди многу побрзо отколку што правото е во состојба да ги нормира. Поради тоа, стандардот de facto станува норма многу порано пред правото нормативно да го поддржи таквиот стандард.

Правната рамка за создавање и употреба на вештачка интелигенција треба да усвои пристап насочен кон човекот, спротивставувајќи се - колку што е можно – на неизбежната дехуманизација на пристапот предизвикана од развојот на технологијата. Треба да се воспостави правна шема која гарантира дека вештачката интелигенција функционира на сигурен и разбирлив начин, не е создадена врз основа на вградена (или наследена) дискриминација и не се користи како инструмент за манипулација.

Од гледна точка на ефикасноста на идната регулатива е критично секој иден правен инструмент да се совпаѓа и да се надоврзува на постоечките правни шеми и стандарди. Дополнително, со оглед на динамиката со која се случува технолошкиот напредок, како и неговата непредвидливост, важно е идните закони за вештачка интелигенција да бидат конструирани и формулирани на начин на кој ќе бидат што е можно пофлексибилни на промените. Предизвикот на создавање механизми за контрола на високоризичните системи за вештачка интелигенција со максимално почитување на човековите права е веќе актуелен. ЕУ има за цел да стане светски лидер во создавањето безбедна средина за употреба на вештачката интелигенција. Постигнувањето на таа цел подразбира и создавање на соодветна етичка и правна рамка за развој и користење на производи и услуги базирани на технологии за вештачка интелигенција. Во последните неколку години, во рамките на ЕУ и Советот на Европа, донесени се голем број документи, препораки, декларации и предлози со цел да се подигне свеста за загрозеноста на човековите права и потребата од нивна заштита преку создавање на соодветна правна рамка заснована на обврзувачки и необврзувачки норми. Соодветен пристап кон заштитата на човековите права од производи и услуги базирани на технологии за вештачка интелигенција е

содржан во Предлог-Регулативата (Законот) на ЕУ за вештачка интелигенција[15]. Затоа, законските решенија вклучени во оваа идна Регулатива на ЕУ секако ќе претставуваат камен-темелник за регулирање на ова прашање во националните рамки на земјите-членки на ЕУ, како и на многу други земји.

ЕУ дефиниција предвидена во ВИ Актот: Вештачката интелигенција е брзо развивачко семејство на технологии што може и веќе придонесува за широк спектар на економски, еколошки, културолошки и општествени придобивки, доколку се развие во согласност со релевантните општи правни и етички принципи во согласност со Повелбата и вредностите на кои е основана Унијата. Интернетот и европскиот Единствен пазар за дигитални услуги без граници бара зголемена соработка меѓу земјите-членки со цел да се гарантира ефективен надзор и спроведување на новите правила утврдени во предложениот Закон. Употребата на вештачка интелигенција во ЕУ ќе биде регулирана со Законот за вештачка интелигенција, првиот сеопфатен закон за вештачка интелигенција во светот.

Како дел од својата дигитална стратегија, ЕУ сака да ја регулира вештачката интелигенција и да обезбеди подобри услови за развој и употреба на оваа иновативна технологија. Вештачката интелигенција може да донесе многу придобивки, како што се подобра здравствена заштита, побезбеден и почист транспорт, поефикасно производство, поевтина и поодржлива енергија. Европската комисија започна да работи на законот пред две години кога вештачката интелигенција само што почнуваше да се појавува како сервис. Денес веќе има неколку сервиси кои се и комерцијално достапни. Најпопуларни се ChatGPT, Dall-E и Midjourney.

Во април 2021 година, Комисијата ја предложи првата регулаторна рамка на ЕУ за вештачка интелигенција. Како дел од неа, се предлага да се анализираат и класифицираат системите со вештачка интелигенција според ризикот што го претставуваат за корисниците. Различни нивоа на ризик ќе значат повеќе или помалку регулирање. Откако ќе бидат одобрени, ова ќе бидат првите правила за вештачка интелигенција во светот. Приоритет на Европскиот парламент е да се погрижи системите за вештачка интелигенција во ЕУ да се безбедни, транспарентни, да можат да се следат, да бидат недискриминаторски и еколошки. Тие треба да бидат надгледувани од луѓе за да се спречат негативни исходи. Парламентот, исто така, сака да воспостави технолошки неутрална и единствена дефиниција за вештачка интелигенција која може да се примени на идните системи за вештачка интелигенција.

Принципите на Организацијата за економска соработка и развој (ОЕЦД) за вештачка интелигенција[16] беа усвоени во мај 2019 година од земјите-членки кои ја одобрија Препораката на Советот на ОЕЦД за вештачка интелигенција. Овие принципи промовираат вештачка интелигенција која е иновативна и доверлива и која ги почитува човековите права и демократските вредности. Принципите за вештачка интелигенција на ОЕЦД се првите такви принципи потпишани од владите на земјите членки. Тие вклучуваат конкретни препораки за јавната политика и стратегија, а нивниот општ опсег гарантира дека тие можат да се применат во развојот на вештачката интелигенција ширум светот. Принципите промовираат инклузивен раст, правичност, вредности насочени кон човекот, транспарентност, безбедност, сигурност и одговорност и градење човечки капацитети и поттикнување на меѓународна соработка. Препораката, исто така, ги охрабрува националните политики и меѓународната соработка за инвестирање во истражување и развој и поддршка на поширокиот дигитален екосистем за вештачка интелигенција.

---

[15] ЕУ ВИ Акт: првата регулатива за вештачка интелигенција. Европски парламент., преглед на 10.12.2023 год.

[16] Преглед на ВИ принципи. ОЕЦД., преглед на 20.12.2023 год.

За ризиците и заканите од ВИ: Иако многу системи за вештачка интелигенција носат минимален ризик, тие треба да се проценат, бидејќи оние системи кои се сметаат за закана за луѓето се неприфатливи и треба да се забранат. Во редот на највисоките ризици спаѓаат: когнитивна бихејвиорална манипулација на корисниците или на одредени ранливи групи (деца, пациенти, работници, студенти, стари лица...); класификација на луѓето врз основа на однесување, социоекономски статус или лични карактеристики; управување со инфраструктура и процеси; пристап и користење на одредени јавни услуги; помош во правното толкување и примена на законите; и биометриски системи за идентификација во реално време и далечинска биометриска системска идентификација, како што е препознавање на лица, освен по исклучок кога идентификацијата се одвива со задоцнување, а се користи врз основа на судско одобрение за гонење на тешки кривични дела[17].

Исто така, вештачката интелигенција има потенцијал да го зајакне авторитарното управување, да управува со смртоносно автономно оружје, да ја формира основата за помоќни алатки за масовна или индивидуална контрола, надзор и цензура, а системите за препознавање лица може да се претворат во масовен надзор на нашите јавни простори, уништувајќи го секој концепт на приватност. Системите за вештачка интелигенција кои се користат во системот на кривичната правда за предвидување на идното криминално однесување веќе се покажа дека ја зајакнуваат дискриминацијата и ги поткопуваат правата, вклучително и пресумпцијата на невиност. Ризиците се најприсутни кај моделите за вештачка интелигенција кои се наменети за широка употреба (биотехнологија и сајбер-безбедноста), но и кај оние кои се користат во исклучително специјализирани полиња и кои би можеле да предизвикаат штета.

Според предлог тесктот на Актот за вештачка интелигенција, алатките со вештачка интелигенција ќе бидат класифицирани во зависност од ризикот. Алатките се поделени во зависност од ризикот во три групи: минимален, висок и неприфатлив. Така, апликациите наменети за општествено рангирање на граѓаните, како што се апликациите за социјално рангирање (во НР Кина), се во групата неприфатливи и тие треба да бидат забранети во целост. Понатаму, во редот на забранети системи на вештачка интелигенција спаѓаат и далечинските биометриски системи за идентификација „во реално време“ во јавно достапни простори, системи за биометриска категоризација кои користат чувствителни карактеристики (на пр. пол, раса, етничка припадност, статус на државјанство, религија, политичка ориентација) [18]. Понатаму, пример за алатка од висок ризик е вештачката интелигенција која се користи за анализа и рангирање на апликанти за работа според нивното резиме (CV), за биометриски надзор или препознавање емоции. Вакви алатки се дозволени, но за да можат да функционираат ќе мора да бидат исполнети строги законски обврски. Сите високоризични системи за вештачка интелигенција треба индивидуално да се проценат и да се регулираат пред да бидат понудени на пазарот и ставени во употреба и треба активно да се следат за време на нивното користење. Системите со ограничен ризик ќе треба да ги исполнат минималните барања за транспарентност што ќе им овозможат на корисниците да донесат информирани одлуки. Пред интеракцијата со апликацијата, корисникот треба да биде информиран и свесен за можностите на самата апликација и евентуалните ризици од нејзиното користење за да може да одлучи дали да продолжи да ја користи. Корисниците треба да знаат кога се во интеракција со вештачката интелигенција, особено во ситуации кога корисниците комуницираат со системи кои генерираат или манипулираат со содржини со слики, аудио или видео податоци (на пр. deepfakes, EMU). И, во тек со најновите трендови, генеративните системи за вештачка интелигенција како ChatGPT ќе мора да се усогласат со барањата за транспарентност и јасно и

[17] Европратениците се подготвени да преговараат за први правила за безбедна и транспарентна вештачка интелигенција. Европски парламент.

[18] ЕУ ВИ Акт: првата регулатива за вештачка интелигенција,. Европски парламент, преглед на 10.12.2023 год.

недвосмислено да објават дека содржината е генерирана од вештачка интелигенција, а корисникот изречно да се согласи со понудените услови за користење на апликацијата.

ЕУ во моментов го подготвува првиот сет на сеопфатни правила во светот за управување со можностите и заканите на вештачката интелигенција. Целта е да се претвори ЕУ во глобален центар за доверлива вештачка интелигенција. Законот за вештачка интелигенција претставува прва регулаторна рамка на ЕУ за вештачка интелигенција. Законот за вештачка интелигенција утврдува јасни одговорности за земјата-членка која ја следи усогласеноста на давателите на услуги основани на нејзината територија со обврските утврдени во предложениот закон[19]. Ова обезбедува најбрза и најефикасна примена на правилата и ги штити сите граѓани на ЕУ. Таа има за цел да обезбеди едноставни и јасни процеси и за граѓаните и за давателите на услуги да најдат олеснување во нивните интеракции со надзорните органи. Кога ќе се појават системски ризици низ Унијата, предложениот закон предвидува надзор и спроведување ширум Унијата. На 14 јуни 2023 година, Европскиот парламент ја усвои својата преговарачка позиција за Законот за вештачка интелигенција и сега се во тек преговори со земјите од ЕУ во Советот за конечната форма и содржина на законот. Целта е да се постигне договор до крајот на 2023 година. Приоритет на Парламентот е да се осигура дека системите за вештачка интелигенција што се користат во ЕУ се безбедни, транспарентни, следливи, недискриминаторски и еколошки. Парламентот, исто така, сака да воспостави технолошки неутрална, единствена дефиниција за вештачката интелигенција која би можела да се примени на идните системи за вештачка интелигенција. На 8 декември 2023 година, преговарачите од Парламентот и Советот постигнаа политички договор со кој истакнуваат дека ги препознаваат потенцијалните ризици за човековите права и демократијата кои може да произлезат од употребата на вештачката интелигенција во определени апликации, па согласни се дека треба да се забрани нејзината употреба за цели на:

- Дизајнирање и употреба на системи за собирање и обработка на биометриски податоци;
- Преземање и користење на фотографии од ликови од системите за масовен надзор без правна основа и креирање на бази на фотографии;
- Утврдување на емотивна состојба на работното место или во образовна институција;
- Социјално рангирање базирано на однесувањето на поединците и нивните лични карактеристики;
- Дизајнирање на алатки врз база на вештачка интелигенција кои манипулираат со однесувањето на поединците;
- Дизајнирање и користење на алатки врз база на вештачка интелигенција со цел користење на ранливоста на поединците (пол, возраст, попреченост, социо-економска состојба) за цели спротивни на етичките норми.

Исклучоци ќе постојат само во строго определени ситуации поврзани со спроведувањето на законите со користење на системи за биометриска идентификација. Овие системи ќе може да бидат користени исклучиво на јавни површини, за цели на спроведување на закон, со претходно издаден судски налог за конкретна цел како што е потрага по лица осомничени за сериозен криминал, потрага по жртви, превенција од тероризам.

За системите кои користат ВИ, а кои ќе бидат класифицирани како високо ризични по правата на граѓаните, договорени се мерките за задолжително спроведување на анализа на влијанието врз човековите права. Граѓаните ќе имаат и право да вложат приговор за неправичното користење на вештачката интелигенција и да добијат објаснување на целите и начините на кои системите се користат.

[19] Акт на ЕУ за вештачка интелигенција: прва регулатива на оваа област, Вербативен извештај на постапката - Закон за вештачка интелигенција (A9-0188/2023). Европски парламент.

3. Дигитални права и човекови права

Потеклото на човековите права произлегува од теоријата на природното право. Човековите права како природни права ги стекнува секое човечко суштество со раѓање. Човековите права претставуваат неотуѓиви права и слободи и подеднакво важат за сите луѓе, без разлика на раса, пол, јазик, религија, економски статус, образование, политичко и друго мислење, во какви било околности. Без оглед на различноста меѓу општествата и луѓето, човековите права ја формираат нишката што ги поврзува заедно. Тие претставуваат универзални вредности, кои се заеднички за сите луѓе. Во суштината на концептот на човековите права е стремежот за заштита на човековото достоинство. Тој ја става личноста на човекот во центарот на вниманието и е заснован на заеднички општ систем на вредности.

Бројни меѓународни и национални правни акти континуирано ја прошируваат листата на заштитени човекови права, така што практично е невозможно да се создаде дефинитивен каталог на човекови права и слободи. Типичен пример за проширување на листата на човекови права и слободи се правните документи на Европската Унија. Тие значително ги проширија човековите права и слободи во областа на заштитата на потрошувачите, заштитата на интелектуалната сопственост, заштитата на природната средина, заштитата на податоците, заштитата на работничките права, а веќе се подготвуваат нови правни акти кои ќе ја прошират листата на човекови права и слободи во областите на заштита од вештачка интелигенција и алгоритамска дискриминација.

Концептот наречен „Интернетот како основно човеково право“ е иницијатива за која се залагаат групи луѓе и поединци со цел да се прогласи Интернетот за основно човеково право во сите земји во светот. Признавањето на Интернетот како човеково право има за цел да обезбеди сите луѓе да имаат еднакви можности и пристап до дигиталниот свет и дека Интернетот се користи како алатка за промовирање на човековата слобода, информации, образование и социјален развој. Признавањето на Интернетот како човеково право подразбира заштита на слободата на изразување на Интернет. Државите треба да гарантираат дека нема да има неоправдани ограничувања на пристапот до информации и слободата на изразување на Интернет, освен во случаи кога таквите ограничувања се строго неопходни и пропорционални за заштита на други основни права или интереси. Конечно, признавањето на Интернетот како човеково право ја става одговорноста на државите да ја заштитат приватноста и податоците на корисниците на Интернет. Ова вклучува усвојување законодавство и политики за да се обезбеди безбедност и заштита на податоците, како и регулирање на пристапот и следењето на Интернет за да се спречи злоупотреба.

Јануари 1983 година се смета за официјален датум на појавувањето на интернетот. Ова е моментот кога новата технологија (Протокол за контрола на пренос/Интернет протокол – TCP/IP) создаде стандард кој им овозможи на различни компјутерски мрежи да комуницираат меѓу себе, во реално време, на голема далечина. Она што беше замислено и создадено како алатка за размена на податоци и информации меѓу научниците и професионалците постепено привлече сè повеќе и повеќе корисници и се прошири во глобални рамки. На почетокот на новиот милениум се појавија платформи и апликации за здружување, комуницирање, промовирање, кои означија почеток на една нова ера во користењето на интернетот: услугата за видео/говорни повици Skype започна со работа во 2003 година, Facebook во 2004 година, Twitter (X) во 2006 година, Instagram во 2010 година, Google во 2011 година, TikTok во 2017 година, ChatGPT на OpenAI во 2018 година. Новата дигитална технологија го претвори светот во глобално село. Човештвото во сите делови на земјината топка стана поврзано, брзината на информациите се зголеми експоненцијално и луѓето започнаа да комуницираат едни со други во реално време на долги растојанија. Интернетот требаше да го зближи, едуцира и да го

демократизира целиот свет. Од друга страна, интернетот со својата сеопфатност и отвореност предизвика внимание и интерес и кај друга категорија корисници со не толку чесни намери и побуди. Така, одредени авторитарни држави и влади го открија потенцијалот на новиот медиум и новата технологија и почнаа да ја користат за свои недемократски цели: да пропагираат и пласираат сопствена верзија на „вистината“, да влијаат на јавното мислење во нивна полза, да ги следат и да ги контролираат неистомислениците или целосно да го затворат пристапот до интернетот за определена категорија корисници. Некои индивидуални корисници, но и групации, исто така почнаа да ја злоупотребуваат технологијата преку говор на омраза, сајбер-мобинг и пласирање дезинформации. Овие деструктивни и незаконити феномени постоеле и порано, пред појавата на интернетот, но сега е многу полесно да се активираат, со едноставно кликнување на копче на тастатурата или со допир на екранот, скриени во безличната анонимност на мрежата.

Гледајќи, пак, од позитивен аспект, неколку земји веќе презедоа чекори за експлицитно признавање на пристапот до Интернет како основно право, сметајќи дека е неопходно граѓаните целосно да ги остварат своите права на слобода на изразување, информации и учество во дигиталната ера. Така, во 2000 година, Естонија стана првата земја во светот што го прогласи пристапот до Интернет за човеково право. Финска го направи тоа во 2010 година, со што пристапот до широкопојасен интернет стана законско право за секој граѓанин. Други земји, како Франција, Грција и Костарика, го признаа пристапот до Интернет како основно право преку различни правни рамки или судски одлуки. Може да се заклучи дека, дури и во земјите каде пристапот до Интернет не е експлицитно признаен како човеково право, се прават напори да се премости дигиталниот јаз и да се обезбеди поширок пристап до Интернет, со оглед на неговото зголемено значење во денешното дигитално општество.

Во 2012 година, Советот за човекови права на Обединетите нации (ОН) воспостави важен фундаментален принцип: човековите права и слободи се применуваат подеднакво онлајн и офлајн, а дигиталните права претставуваат човекови права. Сите луѓе имаат право слободно да пристапуваат, користат, создаваат и објавуваат информации, да ја користат слободата на изразување, споделување информации и комуникации без при тоа да ги нарушуваат правата на другите. Се гарантира правото на секој човек да пристапува, користи, создава, споделува и објавува информации преку дигитални медиуми, блогови, веб-страници и слично, сè додека се почитуваат правата на другите. Сепак, предизвик е овие права да се реализираат подеднакво за сите граѓани и да се поттикнат државите да создадат поволна средина за дигитални инвестиции и развој, но и за унапредување на демократските вредности и човековите права. Владите, односно националните институции за човекови права во соработка со граѓанските организации, треба да се стремат да ги обезбедат правата на корисниците на дигиталните услуги и технологии и да ги заштитат истите од злоупотреби, пред сè на регулаторно и институционално национално ниво, но и со учество во креирање на меѓународни правни акти (декларации, директиви, повелби, препораки) и активности. Во 2016 година, Советот за човекови права на Обединетите нации донесе необврзувачка резолуција (soft law instrument) со која се осудува намерното попречување на пристапот до интернет од страна на владите, нагласувајќи ја при тоа важноста на слободата на изразување на интернет.

Исто така, право на ефективна правна заштита за прекршување на човековите права и основните слободи треба да се добие директно и од давателите на услуги на Интернет, при што корисникот може, во зависност од повредата, да достави приговор, жалба, барање, и да добие објаснување, одговор, исправка, извинување, враќање на работа, повторно поврзување или компензација за сторениот прекршок и повреда и за причинетата штета. Давателите на пристап до онлајн содржини и услуги треба да ги направат информации за правата на корисниците на дигиталните услуги лесно достапни и видливи, а државата има обврска да ги заштити корисниците,

нормативно и административно, од противзаконски активности извршени при користење на Интернет и други дигитални содржини.

Советот на Европа во 2014 година издаде Водич за човекови права за корисниците на Интернет[20] со цел да помогне корисникот подобро да ги разбере правата што ги ужива во онлајн опкружувањето и да даде насоки што може да се направи кога правата се загрозени или повредени. Водичот ги објаснува, на едноставен и разбирлив јазик, правата и слободите содржани во Европската конвенција за човекови права, меѓународен договор што ги обврзува 47-те земји-членки на Советот на Европа, и како тие се применуваат на Интернет. Овој Водич е наменет за: помагање на поединци кои се соочуваат со потешкотии во остварувањето на нивните права; да им помогне на владите и институциите да ги исполнат своите обврски за заштита и почитување на човековите права; промоција на човековите права на корисниците на интернет; промовирање на корпоративната општествена одговорност преку охрабрување на приватниот сектор да дејствува одговорно и со почитување на човековите права на поединците со кои стапува во договорни односи. Водичот е документ кој се развива и треба периодично да се ажурира и да се доработува паралелно со појавувањето на нови технолошки трендови.

Во дигиталното опкружување, според овој Водич, на секој корисник на дигиталните услуги треба да му бидат овозможени и загарантирани следните права:

- **Пристап до интернет и недискриминација** - ниту еден корисник не треба да биде исклучен од Интернет против негова волја, освен кога за тоа ќе одлучи суд (овде не станува збор за исклучување по вина на корисникот поради неисполнување на договорните обврски кон давателот на услуги). Покрај тоа, пристапот до Интернет треба да биде неселективен и недискриминаторски.
- **Слобода на изразување и информации** - секој корисник може слободно да се изразува онлајн и да споделува или да пристапува до информации и мислења, вклучително и оние што можат да навредат или вознемируваат, притоа почитувајќи ја туѓата репутација и приватност. Државите имаат должност да го почитуваат и заштитат ова право. Секое ограничување на слободата на изразување мора да има легитимна цел во согласност со Европската конвенција за човекови права и домицилната регулатива, па така ограничувањата може да важат за објави кои поттикнуваат дискриминација, омраза или насилство или кои претставуваат закана за националната безбедност или јавниот ред и мир. Корисниците можат да одберат да не го откриваат својот идентитет на интернет, но треба да знаат и да бидат информирани дека националните власти можат да преземат мерки што може да доведат до откривање на идентитет, во строго определени случаи и по исклучок.
- **Слобода на собирање, здружување и учество** - секој корисник има слобода да користи која било веб-локација, апликација или друга услуга за да оствари контакт, да соработува и да се дружи со своите истомисленици, пријатели, колеги, соработници. Исто така, овозможено и загарантирано е и правото на мирен онлајн протест. Сепак, корисникот треба да биде информиран дека може да се соочи со правни последици доколку таквиот протест и онлајн здружување доведе до блокирање, прекин на услугите или оштетување на имотот или загрозување и повреда на правата на други луѓе.
- **Приватност и заштита на личните податоци** - личните податоци треба да се обработуваат само со изречна согласност на корисникот или ако тоа е пропишано со закон. Корисникот на дигиталните услуги секогаш треба да биде информиран дали и на кој начин неговите лични податоци се обработуваат, чуваат или се пренесуваат на други страни и кога, од кого и за каква цел. Исто така, обврска и право на секој корисник е да се грижи за своите лични податоци и да врши контрола врз нив (да ја провери точноста,

[20] Водич за човекови права за корисници на Интернет. Совет на Европа., преглед на 22.11.2023 год.

да побара корекција или бришење). Забрането е било кој корисник да биде предмет на општ надзор, снимање или прислушување, освен во исклучителни околности пропишани со закон, заради кривична истрага или постоење на основан сомнеж или доказ за загрозување на безбедноста и јавниот ред и мир, врз основа на посебно судско овластување.

- **Заштита на деца и млади луѓе** - децата и младите имаат право на посебна заштита при користење на интернет. Поради недостаток на свест и разбирање за технологијата и онлајн светот, децата се особено ранливи на прекршување на приватноста. Овие прекршувања на приватноста може да вклучуваат откривање на лични и/или чувствителни информации на детето, како и објавување на лажни или погрешни информации што го прикажуваат детето во негативно светло, било со погрешно прикажување на неговите верувања, постапки или карактер. Дигиталниот простор е моќна алатка, но исто така може да ја загрози физичката и менталната благосостојба на децата. Играчките, компјутерските или игрите на таблет или мобилен телефон и другите детски производи често користат технологија за вештачка интелигенција, која може да собира огромни количини лични податоци од децата без нивно знаење или согласност. Бидејќи децата сè повеќе се потпираат на вештачката интелигенција, тие може да развијат лажно чувство на сигурност, што ќе ги наведе да веруваат дека технологијата е доверлива и безбедна. Кога станува збор за безбедноста на децата, вештачката интелигенција може да ги изложи на потенцијално штетна онлајн содржина и вознемирувачки искуства, како што се сајбер малтретирање, говор на омраза и изложеност на графичко насилство или експлицитни содржини. Понатаму, алгоритмите за вештачка интелигенција може да се користат за ширење лажни информации или за манипулирање со децата со цел да се вклучат во ризични активности како компјутерски криминал или самоповредување. Децата можеби не се свесни за ризиците од споделување лични информации на интернет или можеби не знаат кога и самите се цел на лажни или штетни информации. Доколку содржината што ја објавиле самите или ја објавиле нивните родители, старатели или други лица од семејството, училиштето или најблиското опкружување, а истата го загрозува нивното достоинство, безбедност, приватност или може да биде штетна за нив во иднина, таа треба да се избрише во најкраток временски период, на барање на корисникот или неговиот старател. Исто така, треба да бидат заштитени од мешање во нивната физичка и ментална состојба, особено во поглед на сексуалната експлоатација и злоупотреба на содржини (податоци, фотографии, видеа) заради нечесни и лукративни цели.

Во редот на меѓународните документи и правни акти за регулирање на вештачката интелигенција спаѓа и Декларацијата за Европските дигитални права и принципи, усвоена во декември 2022 година од страна на Европската комисија[21]. Од новата Декларација се очекува да ги надополни постојните правни акти кои ги регулираат заштитата на податоците и е-приватноста, како и Повелбата за фундаментални права, и да обезбеди насоки за ЕУ и земјите-членки во процесот на прилагодување на дигиталната трансформација. Технологијата треба да им служи и да им користи на сите луѓе кои живеат во ЕУ и да ги поттикне да ги остварат своите аспирации, но при тоа не треба да ја нарушува нивната безбедност или да ги загрозува основните права.

Принципите содржани во Декларацијата се обликувани околу следните главни теми:

1. Ставање на луѓето и нивните права во центарот на дигиталната трансформација.
2. Солидарност и инклузија - Секој човек треба да има пристап до технологијата, која треба да биде инклузивна и да ги промовира правата на луѓето.

[21] Европски дигитални права и принципи. Европска комисија., преглед на 18.12.2023 год.

3. Обезбедување слобода на избор на интернет - Секој човек треба да биде овластен да направи свој, информиран избор на интернет, вклучително и при интеракција со вештачка интелигенција и алгоритми кои мора да бидат транспарентни.
4. Слобода на избор - Слобода да избираме кои онлајн услуги ги користиме, врз основа на објективни, транспарентни и веродостојни информации.
5. Поттикнување на учество во дигиталниот јавен простор - Конкретно, потписниците ќе се обврзат да дејствуваат во повеќе области, вклучувајќи: поврзување; дигитално образование, обука и вештини; фер и правични услови за работа; дигитални јавни услуги, плуралистичка јавна дебата и учество во демократијата.
6. Зголемување на безбедноста и сигурноста - Секој треба да има пристап до безбедни дигитални технологии, производи и услуги кои ја штитат приватноста. Дигиталните принципи се обврзуваат да ги заштитат интересите на луѓето, бизнисите и јавните услуги од сајбер криминалот и да се спротивстават на оние кои се обидуваат да ја нарушат безбедноста и интегритетот на нашата онлајн околина.
7. Промовирање на одржливоста на дигиталната иднина - Дигиталната и зелената транзиција се тесно поврзани. Иако дигиталните технологии нудат многу решенија за климатските промени, мора да се внимава тие самите да не придонесуваат за проблемот. Дигиталните производи и услуги треба да бидат дизајнирани, произведени и отстранети на начин што нивното користење ќе го намали нивното влијание врз животната средина и општеството. Исто така, треба да има повеќе информации во врска со влијанието врз животната средина и потрошувачката на енергија кај таквите услуги.

Еден од начините за ефикасна борба против негативното влијание на вештачката интелигенција врз основните права и слободи би можело да биде воспоставувањето обврски за креаторите и имплементаторите на системи за вештачка интелигенција да обезбедат почитување на човековите права и владеењето на правото при создавањето на таквите системи (човекови права по дизајн) и спроведување соодветни технички и организациски мерки при експлоатација на системот за вештачка интелигенција (човекови права според модел/образец). Членот 25 од Општата регулатива за заштита на податоците (GDPR) во ЕУ воведе два нови термини: приватност по дизајн и приватност по образец. Приватноста по дизајн подразбира почитување и имплементирање на сите принципи за заштита на приватноста уште во најраната фаза на дизајнирање системи за собирање и обработка на податоци за проактивно и превентивно дејствување. Приватноста по дизајн подразбира примена на соодветни технички и организациски мерки од страна на ракувачите со податоци, со цел да се обезбеди усогласеност со сите принципи за заштита на податоците при собирањето и обработката на податоците.

Слично на решението во Општата регулатива за заштита на податоците на ЕУ во однос на усогласеноста со правото на приватноста по дизајн, може да се воведе законска обврска за почитување на човековите права и владеењето на правото по дизајн и по образец за сите креатори и спроведувачи на системи за вештачка интелигенција кои можат да ги загрозат основните права и слободи на сите поединци. Оваа обврска би се состоела во проактивно и превентивно вградување на основните принципи за заштита на човековите права при создавањето на системи за вештачка интелигенција, како и обврска на оние кои управуваат и користат системи за вештачка интелигенција да ги преземат сите неопходни технички и организациски мерки во текот на експлоатација на системи за вештачка интелигенција, со цел да се заштитат основните права и слободи на поединците и владеењето на правото.

4. Алгоритамска дискриминација – автоматизирано носење одлуки и повреди на човековите права

Начелото на еднаквост е едно од основните начела на човековите права, а од него произлегува принципот на недискриминација, односно забрана на дискриминација меѓу луѓето. Самиот термин „дискриминација“ потекнува од латинскиот јазик и значи дистинкција, но, барем во правото, овој термин ја изгубил својата неутралност и го добил негативното значење на незаконска дистинкција. Во областа на човековите права, дискриминацијата е дистинкција во однос на поседувањето и обемот на правата, што не е дозволено поради основата и начинот на дистинкција. Забраната за дискриминација во современото право се однесува на дискриминација врз основа на раса, боја на кожа, пол, јазик, религија, политичко или друго уверување, национална и социјална положба, потекло, имот, раѓање и друг статус. Набројувањето не е исцрпено, односно мора да се земат предвид и други слични основи на дискриминација. Едно од најчесто пријавени негативни влијанија на вештачката интелигенција врз човековите права е влијанието врз забраната на дискриминацијата, односно врз правото на еднаков третман.

Алгоритамскиот систем за донесување одлуки може да се дефинира како компјутерски процес кој донесува одлуки самостојно или го поддржува човечкото одлучување. Понекогаш алгоритмот одлучува на сосема автоматски начин. На пример, филтерот за спам за е-пошта може целосно автоматски да филтрира спам пораки од сандачето на корисникот. Понекогаш луѓето донесуваат одлуки со помош на алгоритам и таквите одлуки се делумно автоматски. На пример, врз основа на проценката на кредитната способност на клиентот од страна на системот за вештачка интелигенција, вработениот во банката може да одлучи дали клиентот може да позајми пари од банката. Неопходно е да се направи разлика помеѓу одлуките донесени врз основа на алгоритам кои се целосно автоматизирани и оние кои се само делумно автоматизирани. Ова е важно поради утврдувањето на одговорноста во случај на прекршување на човековите права и поради начинот на кој може да се направат промени во системот за автоматско одлучување, за да не се случуваат повреди на човековите права во иднина. Во алгоритамските системи за донесување одлуки со делумно човечко учество, постои тенденција да се минимизира нечија одговорност со едноставно следење на препораките на компјутерот. Овој феномен се нарекува автоматска пристрасност. Системите за вештачка интелигенција засновани на пристрасни информации може да предизвикаат алгоритамска дискриминација, односно дискриминаторски алгоритамски одлуки или однесувања. Ако системот за вештачка интелигенција учи врз основа на претходни податоци базирани на дискриминаторски одлуки, тогаш тој самиот може да носи дискриминаторски одлуки врз основа на „фидбек циклуси“, односно може да ги загрози човековите права. Последица на алгоритамското управување со работните процеси може да биде дехуманизација и загрозување на правата на вработените. Така, ако во минатото за одредени работни места биле ангажирани повеќе мажи отколку жени, тогаш системот за вештачка интелигенција, врз основа на учење од претходните примери, исто така ќе ги дискриминира жените во иднина при вработување на тие работни места.

Системите на вештачка интелигенција можат да произведуваат различни форми на дискриминација благодарение на фактот што учат од претходните лоши примери или поради тоа што се програмирани да ги стават поединците во дискриминаторска положба. За да се елиминираат овие форми на дискриминација, неопходно е да се корегираат системите за вештачка интелигенција врз основа на примери на идеално однесување без дискриминација. Стереотипните однесувања кои водат до дискриминација не треба да се повторуваат од системите за вештачка интелигенција, но мора да се корегираат за да се укинат неправдите и да се почитуваат правилата за еднаквост за сите. Сите случаи во кои функционирањето на алгоритмот доведува до дискриминација, по која било основа, мора ефективно да се санкционираат.

Обврзувачките и необврзувачките правни норми се главните пречки за алгоритамската дискриминација. Пред сè, постојат обврзувачки прописи на меѓународно и национално ниво за недискриминација и заштита на податоците, но и многу други прописи, стандарди и правила на однесување. Универзалната декларација за човекови права на ОН од 1948[22] година во својот прв член гарантира слобода и еднакви права на сите луѓе, а во членот 2 прокламира и гарантира недискриминација: „Секој има право на сите права и слободи прогласени во оваа Декларација без никаква разлика во однос на расата, бојата, полот, јазикот, религијата, политичкото или друго мислење, националното или социјалното потекло, имотот, раѓањето или други околности". Европската конвенција за заштита на човековите права и основни слободи, во согласност со Универзалната декларација за човекови права на ОН, забранува дискриминација во членот 14 со следната одредба: „Уживањето на правата и слободите наведени во оваа Конвенција ќе се обезбеди без дискриминација по која било основа, како што се пол, раса, боја, јазик, религија, политичко или друго мислење, национално или социјално потекло, здружување со национално малцинство, имот, раѓање или друг статус". Директивата на ЕУ од 2000 година за спроведување на принципот на еднаков третман меѓу лицата без разлика на расното или етничкото потекло[23] признава две форми на дискриминација: директна и индиректна дискриминација. Во членот 2 од оваа Директива: „Директна дискриминација постои кога едно лице се третира понеповолно од друго лице, било третирано или би било постапувано понеповолно во слична ситуација врз основа на расно или етничко потекло". Во истиот член, индиректната дискриминација е дефинирана како што следува: „Индиректната дискриминација се смета дека се јавува кога навидум неутрална одредба, критериум или практика би ги ставила лицата од одредено расно или етничко потекло во особено неповолна положба во однос на други лица, освен ако таа одредба, критериум или практика објективно оправдана со легитимна цел, а средствата за постигнување на таа цел се соодветни и неопходни". Во случај на алгоритамска дискриминација, често се јавува индиректна дискриминација, така што една навидум неутрална одредба доведува до особено неповолна положба за одредена група луѓе. Според тоа, не е релевантно дали дискриминаторот имал намера да дискриминира, туку релевантен е ефектот што алгоритамското одлучување го имало во пракса. Индиректната дискриминација исто така е тешко да се открие поради нетранспарентноста на системот за вештачка интелигенција. Така, на пример, корисниците на услугите на банката често немаат информации дека одлуките за нивните барања за кредит ги носат системи за вештачка интелигенција, а не луѓе. Второ, корисниците на услугите на банката, дури и кога знаат дека за нивните барања одлучува систем за вештачка интелигенција, немаат соодветно објаснување зошто нивното барање е одбиено, па логично е дека клиентите тешко можат да утврдат дали алгоритамската одлука е дискриминаторска или не. Може да се заклучи дека забраната за директна и индиректна дискриминација секако се однесува на директна и индиректна алгоритамска дискриминација, но дека оние кои сакаат да го откријат и докажат овој вид дискриминација се соочуваат со бројни проблеми. Како и во многу други случаи, постојат правни правила, но нивната имплементација бара бројни појаснувања, судска пракса и воспоставување нови стандарди во областа на креирање и примена на системи за вештачка интелигенција.

И Општата регулатива за заштита на податоци (GDPR) многу сериозно го разработува проблемот на алгоритамска дискриминација. Во неа се инсистира на подигање на нивото на транспарентност во сите случаи на обработка на податоци, особено кога станува збор за автоматско одлучување од страна на системите за вештачка интелигенција, што е наведено и во член 13, точка ф. Во таа точка се наведува дека на поединците ќе им се дадат информации за тоа по која методологија и логика се донесувале одлуките, какво е значењето и кои се последиците од таквата обработка на податоци и носење такви одлуки. Точката 71 од Преамбулата, како и членот 22 директно го опфаќаат алгоритамското одлучување. Овде се утврдува правото на поединецот да бара врз него да не се

[22] Универзална декларација за човекови права - англиски. Обединети нации., преглед на 18.12.2023 год.
[23] Директива на Советот 2000/78/ЕЗ од 27 ноември 2000 година за воспоставување општа рамка за еднаков третман во вработувањето и професијата. EUR-Lex., преглед на 22.12.2023 год.

однесува или да нема влијание одлука донесена исклучиво врз основа на автоматизирана обработка на податоци што произведува правни последици или значително го погодува. Наведени се примери со практики за регрутирање преку Интернет. Таквата автоматизирана обработка на податоци вклучува создавање профил, т.е. проценка на личните карактеристики на поединецот, особено оние поврзани со работните резултати, економскиот статус, здравјето, личните преференци или интереси, сигурноста или однесувањето, локацијата или движењето, кога тоа создава правни последици поврзани со поединецот или директно го засегнува. Сепак, одлучувањето засновано на таква обработка, вклучително и профилирање, може да биде дозволено доколку тоа го дозволува правото на ЕУ или правото на земјата-членка на која е предмет обработувачот на податоците, меѓу другото, за целите на следење и спречување кривично дело, измама или даночно затајување, во согласност со прописите, стандардите и препораките на институциите на ЕУ или националните власти. Во секој случај, за таквата обработка треба да се применат соодветни заштитни мерки, кои треба да вклучуваат обезбедување одредени информации на поединецот и право на човечко учество во процесот на одлучување, право да се изрази сопственото мислење, да се добие објаснување на одлуката донесена по ваквата оцена и правото на оспорување на одлуката.

5. Развој на вештачката интелигенција и заштита на личните податоци

Кога станува збор за поврзаноста помеѓу вештачката интелигенција и заштитата на личните податоци, може да се каже дека тоа се две области кои сè повеќе меѓусебно се пресретнуваат. Технологијата на развој на вештачката интелигенција е сè понапредна, а со тоа покрева и бројни прашања кои се однесуваат на етичкото користење на вештачката интелигенција во контекст на заштита на личните податоци на потенцијалните корисници на системите кои користат вештачка интелигенција.

Сите институции, организации, компании кои се вклучени во процесот на развој и користење на вештачката интелигенција мора да обезбедат транспарентност во користењето на вештачката интелигенција, односно, да дадат јасни и концизни информации на корисниците за тоа дали и на кој начин нивните лични податоци се собираат и обработуваат при корисењето на системот кој користи вештачка интелигенција. Дополнително, на корисниците треба јасно да им биде дадено до знаење за која специфична цел нивните лични податоци ќе бидат собрани, складирани или на кој било друг начин обработувани, и да им биде понудена експлицитна можност да ја дадат или да не ја дадат својата согласност.

Прашање кое е од исклучителна важност кога станува збор за поврзаноста на вештачката интелигенција и заштитата на личните податоци е и прашањето на автоматско носење на одлуки. Согласно Општата регулатива за заштита на личните податоци (General Data Protection Regulation-GDPR), секоја индивидуа има право да не биде предмет на автоматско носење на одлуки, вклучувајќи го тука и профилирањето[24]. Ова значи дека во процесот на развој на системи кои користат вештачка интелигенција мора да биде направена анализа на потенцијалното влијание кое би го имало користењето на вештачката интелигенција врз заштитата на личните податоци и дека нема да се носат одлуки за корисниците ниту пак ќе се врши профилирање без да се земат предвид човековите права.

Воспоставувањето на баланс помеѓу потребата за технолошки достигнувања, развој на вештачката интелигенција и потребата за почитување на правилата, насоките и законите за заштита на личните податоци е истовремено и клучна и комплексна активност. Заштитата на личните податоци не смее да биде сфатена само како законска обврска со која мора да се усогласат оние кои работат на развивање на вештачката интелигенција туку како една од основните компоненти за разумно користење на вештачката интелигенција.

Од друга страна пак, заштитата на личните податоци не смее да биде причина за забавување на иновативните процеси и развојот на вештачката интелигенција. На заштитата на личните податоци треба да се гледа како нешто кое ќе обезбеди етичко, фер користење на вештачката интелигенција со почитување на правата на корисниците.

Со цел усогласување со европските регулативи, Република Северна Македонија донесе нов Закон за заштита на личните податоци[25]. Новиот Закон за заштита на личните податоци е во целост е усогласен со Општата регулатива за заштита на личните податоци (GDPR), а

[24] „Профилирање“ е секоја форма на автоматска обработка на лични податоци, која се состои од користење на лични податоци за оценување на одредени лични аспекти поврзани со физичкото лице, а особено за анализа или предвидување на аспекти кои се однесуваат на извршување на професионалните обврски на тоа физичко лице, неговата економска состојба, здравје, лични преференции, интереси, доверливост, однесување, локација или движење

[25] Закон за заштита на личните податоци, Службен весник на РСМ бр.42/20 и бр.294/21

дополнително Агенцијата за заштита на личните податоци[26] има донесено и подзаконски акти за проценка на влијанието врз заштитата на личните податоци.

Трендот на дигитализација на јавните услуги ја наметна и потребата за поддршка на процесот при самото планирање на дигитализацијата, особено од аспект на проценка на можните ризици по приватноста на граѓаните. За таа цел, Фондацијата за интернет и општество – Метаморфозис, во октомври 2023 креираше и објави Методологија за проценка на влијанието на заштитата на личните податоци[27] со која се опишува постапката за вршење на проценката на јавните услуги што се во процес на дигитализација во сите јавни институции, како и проценка на влијанието што вештачката интелигенција ќе го има врз приватноста на граѓаните, доколку се применува од институциите во процесот на испорака на јавните услуги.

Оваа Методологија го опишува методот и ги определува чекорите при спроведување на проценката за влијанието на заштитата на личните податоци (ПВЗЛП) и ги обезбедува потребните критериуми за процена и референтни примери.

Корисници на овој документ се офицерот за заштита на личните податоци и одговорните лица на организациските единици во институцијата.

5.1. Вештачката интелигенција и начелата за заштита на личните податоци

Законот за заштита на личните податоци дефинира дека „личен податок“ е секоја информација која се однесува на идентификувано физичко лице или физичко лице кое може да се идентификува (субјект на лични податоци), а физичко лице кое може да се идентификува е лице чиј идентитет може да се утврди директно или индиректно, посебно врз основа на идентификатор како што се име и презиме, матичен број на граѓанинот, податоци за локација, идентификатор преку интернет, или врз основа на едно или повеќе обележја специфични за неговиот физички, физиолошки, генетски, ментален, економски, културен или социјален идентитет на тоа физичко лице.

Во оваа смисла, сите лични податоци кои корисниците на системите за вештачка интелигенција ги откриваат при пристапувањето и користењето на системот, се сметаат за лични податоци во оној момент кога со нив ќе може да се открие идентитетот на корисникот.

Институциите, организациите, компаниите кои ги развиваат системите за вештачка интелигенција ги утврдуваат целите и начинот на обработка на личните податоци и се јавуваат во улога контролори[28] на тие лични податоци, па оттука се и одговорни за почитување на начелата за заштита на личните податоци.

[26] Агенција за заштита на лични податоци

[27] Методологија за проценка на влијанието на заштитата на личните податоци. Фондација Метаморфозис., преглед на 27.12.2023 год.

[28] „Контролор“ е физичко или правно лице, орган на државната власт, државен орган или правно лице основано од државата за вршење на јавни овластувања, агенција или друго тело, кое самостојно или заедно со други ги утврдува целите и начинот на обработка на личните податоци, а кога целите и начинот на обработка на личните податоци се утврдени со закон, со истиот закон се определуваат контролорот или посебните критериуми за негово определување

- **Начело на „законитост, правичност и транспарентност“**

Согласно првото начело, личните податоци може да се обработуваат доколку за тоа има законска, или друга правна основа, во доволна мера и на транспарентен начин во однос на субјектот на личните податоци.

За да се дефинира основата за обработка на личните податоци, клучно е да се утврди дали алатката или апликацијата која користи вештачка интелигенција служи за остварување на некое законски гарантирано право на корисниците или пак станува збор за услуга или активност која како правна основа ја има согласноста на корисникот.

Пример за обработка на лични податоци врз основа на закон и согласност на субјектот е неодамна промовираниот дигитален асистент – АДА[29], наменета да обезбеди информации за граѓаните и заинтересираните потенцијални странски инвеститори за инвестициските механизми.

Контролори на личните податоци во овој случај се институциите чии услуги/информации се достапни на АДА но и компанијата која ја развила АДА и која се грижи за функционалноста на оваа дигитална алатка.

Самото отпочнување на комуникацијата со АДА се смета за согласност од страна на корисникот да ги сподели податоците за локацијата и IP адресата, во овој случај правна основа за обработка на овие податоци е согласноста. Понатаму, доколку корисникот сака со помош на АДА да оствари некое друго свое законски гарантирано право, податоците кои се потребни за да тоа му биде обезбедено ќе се обработуваат согласно законски пропишаните процедури и ќе се смета дека постои законска основа за конкретната обработка.

До погоре наведените заклучоци е дојдено по пат на детален преглед на веб страницата на АДА иако согласно првото начело, транспарентноста е првата обврска која треба да се исполни од страна на контролорите и сите информации за АДА би требало да се најдат во детално разработена Политика за приватност која ќе биде првото нешто кое секој корисник ќе може да го види пред воопшто да започне да ја користи АДА.

Транспарентноста на обработката на личните податоци подразбира концизност, пристапност и разбирливост. На локацијата на која е достапна самата алатка или апликација која користи вештачка интелигенција, задолжително треба да има конкретни информации до кои корисникот ќе може лесно да пристапи пред да започне да ја користи и сето тоа напишано на лесно разбирлив јазик, а онаму каде е потребно и поткрепено со визуелизации.

Во студијата за „Влијанието на Општата регулатива за заштита на личните податоци врз вештачката интелигенција“[30] објавена од Европскиот парламент во 2020 година, се разликуваат два концепти на обезбедување на транспарентност.

Првиот концепт е поврзан со обезбедување на отворена информација и демонстрирање на отчетност од страна на оние кои ги развиваат и ги поседуваат апликациите и алатките кои користат вештачка интелигенција, а за чие користење е неопходно да бидат откриени лични податоци на корисниците. Во овој случај, информацијата која треба да биде достапна на

[29] АДА - Првиот дигитален асистент во Владата на Република Северна Македонија

[30] Влијанието на вештачката интелигенција врз Општата регулатива за заштита на податоците. Единица за научно предвидување на службата за истражување на Европскиот парламент.

корисниците, а која се однесува на заштитата на личните податоци треба да вклучува податоци за тоа:

- кој ги обработува личните податоци, односно, кој е контролор на збирката на лични податоци која се креира при користењето;
- кој е основот за обработка на личните податоци, односно, дали личните податоци се собираат согласно закон или пак согласноста на корисникот е единствениот основ за обработката;
- за која цел се собираат личните податоци, односно, што корисникот добива од алатката, услугата, апликацијата;
- кои категории на лични податоци се обработуваат, односно, опис на сите лични податоци кои се собираат;
- колку долго се чуваат личните податоци, односно, кој е рокот во кој личните податоци ќе бидат обработувани во врска со конкретната цел;
- дали се врши пренос на личните податоци во друга земја и која е таа земја;
- дали се врши автоматско донесување на одлуки или профилирање, доколку се врши, која е целта за тоа;
- дали личните податоци се откриваат на трети страни, доколку се откриваат, кои се третите страни;
- кои се правата на корисниците за заштита на нивните лични податоци и како можат тие права да се остварат.

Вториот концепт се однесува на транспарентноста во случаите кога личните податоци на корисниците, освен за првично дефинираната цел, се користат и за унапредување на функционалностите на системот и намалувањето на можностите за грешки. Во овие случаи, информацијата која треба корисниците да ја добијат, треба да вклучува податоци за:

- кои лични податоци ќе бидат предмет на дополнителна обработка;
- на кој начин ќе се врши профилирањето за да се оствари целта;
- на кој начин податоците ќе бидат заштитени при вршењето на дополнителната обработка;
- што презема контролорот за да обезбеди заштита на правата на корисниците од потенцијални ризици поврзани со дискриминација врз основа на расно или етничко потекло, политички ставови, верски или филозофски убедувања или членство во синдикални организации, како и генетски податоци, биометриски податоци, податоци што се однесуваат на здравјето или податоци за сексуалниот живот или сексуалната ориентација.

- **Начело на „ограничување на целите“**

Согласно начелото на „ограничување на целите“, личните податоци се собираат за конкретни, јасни и легитимни цели и нема да се обработуваат на начин што не е во согласност со тие цели. Натамошната обработка за цели на архивирање од јавен интерес, за научни или историски истражувања или за статистички цели, нема да се смета дека не е во согласност со првичните цели за кои се собрани личните податоци.

Клучно за начелото на „ограничување на целите“ е да се разбере дека тоа е тесно поврзано со правната основа за обработка на личните податоци, особено кога правна основа е согласноста на корисникот. Кога ја дава својата согласност за обработка на личните податоци, корисникот треба да биде информиран за целта или целите за кои тие податоци ќе бидат обработувани.

Секоја потреба за дополнителна обработка на личните податоци треба да биде предмет на анализа од страна на контролорот за да може да се утврди дали станува збор за сосем нова цел или пак дополнителна обработка поврзана со постоечката цел за обработка.

Во секој случај, новата обработка ја наметнува обврската од дополнително информирање на корисниците, а онаму каде е потребно и обезбедување на дополнителна согласност.

- **Начело на „минимален обем на податоци“**

Начелото на „минимален обем на податоци“ предвидува дека личните податоци кои се обработуваат за конкретна цел треба да се соодветни и релевантни за конкретната цел и ограничени на она што е неопходно во однос на целта заради која се обработуваат.

Почитувањето на ова начело во контекст на развој и користење на вештачката интелигенција е често предизвик, особено во случаите кога станува збор за користењето на големите бази на податоци за цели на аналитика со помош на вештачка интелигенција и различни статистички методи. Во обид да се надмине овој предизвик, контролорите треба да разберат дека „минималниот обем на податоци“ не ја исклучува можноста да се соберат и обработуваат дополнителни лични податоци но се додека тоа е од корист и за корисникот и не претставува ризик по приватноста.

Унапредувањето на системите за вештачка интелигенција не смее да биде оневозможено од тесното толкување на начелото на „минимален обем на податоци“, но исто така не смее и да се случува без да се постави одреден баланс со потребата за заштита на приватноста на корисниците. Она што контролорите треба да го направат е или да ги псевдонимизираат[31] личните податоци кои сакаат да ги користат и во иднина за цели на анализа или пак да креираат сетови на статистички податоци кои во ниту една ситуација нема да го откриваат идентитетот на корисниците. Со примената на овие механизми, може да се обезбедат саканите разултати кои ќе се однесуваат на група на корисници но можноста да се издвои индивидуа од таа група нема да постои, а со тоа и ризикот да се наруши приватноста на корисникот ќе се сведе на минимум.

- **Начело на „точност“**

Согласно начелото на „точност“, личните податоци треба да се точни и каде што е потребно ажурирани, при што ќе се преземат сите соодветни мерки за навремено бришење или коригирање на податоците што се неточни или нецелосни, имајќи ги предвид целите заради кои биле обработени.

Во контекст на развој и користење на вештачка интелигенција, ова начело е особено важно, особено во ситуациите кога со помош на вештачката интелигенција се прави некаква оценка за корисниците, им се даваат насоки или се носи одлука за нив. Користењето на неточни лични податоци може да биде штетно не само по правото на приватност и заштита на личните податоци, туку и по други права кои се повразни со алатката или апликацијата која се користи. Ова особено може да биде голем предизвик кога се обработуваат и некои од посебните

[31] „Псевдонимизација“ е обработка на личните податоци на таков начин што личните податоци не можат повеќе да се поврзат со одреден субјект на лични податоци без да се користат дополнителни информации, под услов таквите дополнителни информации да се чуваат одделно и да подлежат на технички и организациски мерки со кои ќе се обезбеди дека личните податоци не се поврзани со идентификувано физичко лице или физичко лице кое може да се идентификува

категории на лични податоци[32], на пример, доколку користиме апликација за предвидување на ризик по здравјето, а дел од личните податоци се неточни или нецелосни, резултатот кој ќе го добиеме може да има штетни последици и по здравјето.

- **Начело на „ограничување на рокот на чување“**

Личните податоци треба да бидат чувани во форма која овозможува идентификација на субјектите на личните податоци, не подолго од она што е потребно за целите поради кои се обработуваат личните податоци. Личните податоци може да се чуваат подолго од нивниот рок на чување ако се обработуваат само за целите на архивирање од јавен интерес, за научни или историски истражувања или за статистички цели, а со применување на соодветни технички и организациски мерки заради заштита на правата и слободите на субјектот на личните податоци.

Ограничувањето на роковите на чување е можеби и најголемиот предизивик кога станува збор за складирање и чување на лични податоци во системите за вештачка интелигенција од причина што обемот на податоци е огромен, а самата обработка е комплексна. Дополнително, чувањето на личните податоци е поврзано и со целите за кои тие податоци се првично собрани и понатаму обработувани. За да обезбедат примена на ова начело, контролорите треба, секаде каде што е применливо, да ги преземат мерките на анонимизација, псевдонимизација и користење на податоците за статистички цели. На овој начин ќе се обезбеди ограничување на рокот на чување на податоци со кои може да се идентификуваат корисниците, а сепак ќе се обезбедат и доволен обем на податоци за вршење на анализи и планирања на унапредување на користењето на вештачката интелигенција.

- **Начело на „интегритет и доверливост“**

Обработката на личните податоци треба да се врши на начин кој обезбедува соодветно ниво на безбедност, вклучувајќи заштита од неовластена или незаконска обработка, како и нивно случајно губење, уништување или оштетување, со примена на соодветни технички или организациски мерки.

За да се обезбеди усогласеност со ова начело, треба да се направи анализа на ризикот врз безбедноста на личните податоци и да се дефинираат технички мерки на заштита кои ќе гарантираат дека пристап до личните податоци имаат единствено лица кои имаат овластување за обработка, има воспоставено механизам за ажурирање и проверка на личните податоци за да се гарантира дека истите се точни и целосни и и ажурирани. Дополнително, контролорите треба да обезбедат континуирана достапност, односно, непречен пристап и расположливост (business continuity) на информацискиот систем/системот за вештачка интелигенција.

5.2. Вештачката интелигенција и автоматското носење на одлуки

Брзиот развој на новите технологии несомнено овозможува автоматското носење на одлуки со помош на користење на вештачката интелигенција и големите бази на лични податоци да биде можно врз основа на оние лични податоци кои се обработуваат во системите за вештачка интелигенција, без вклученост на корисникот во времето кога таа одлука се носи.

[32] „Посебни категории на лични податоци“ се лични податоци кои откриваат расно или етничко потекло, политички ставови, верски или филозофски убедувања или членство во синдикални организации, како и генетски податоци, биометриски податоци, податоци што се однесуваат на здравјето или податоци за сексуалниот живот или сексуалната ориентација на физичкото лице.

Еден аргумент во корист на оние кои развиваат системи за вештачка интелигенција е тоа дека овие системи можат да ја избегнат грешката која луѓето се склони да ја прават заради предрасудите кои ги имаат, па системот за вештачка интелигенција нема да исклучи или дискриминира никого врз основа на етничка припадност, пол, припадност на политичка партија, сексуална ориентација и слично. Исто така, автоматското донесување на одлуки со користетње на вештачка интелигенција во процеси како што се на пример инвестиции, вработување, кредитна способност, многу често се покажува како попрецизно отколку кога тие одлуки би ги носеле само луѓето.

Овие аргументи се валидни само во случај кога развојот и користењето на вештачката интелигенција се базираат на претходна анализа на ризикот, собирање на соодветен обем на податоци и поставување на јасни критериуми за анализа, тестирање на системот и примена на соодветни технички мерки на заштита.

При автоматското носење на одлуки со користење на вештачка интелигенција секогаш мора да се тргне од принципот на правичност, кој меѓудругото подразбира дека при носењето на одлуки не треба да бидат вклучени чувствителните лични податоци. На овој начин ќе се избегне дискриминација на индивидуите по која било основа. Особено внимание треба да се посвети во процесот на дефинирање на алгоритмите за да може да се вгради принципот на правичност при автоматското носење на одлуки.

На пример, компанија одлучува да врши селекција на кандидати за вработување со помош на вештачка интелгенција. Инструкциите кои треба да бидат дадени во процесот на развој на вештачката интелигенција се од клучно значење и истите треба да бидат концизни и да ги вклучат само оние лични податоци кои се поврзани со квалификациите за конкретното работно место. Доколку претходно во таа компанија тимот за човечки ресурси имал пракса да исклучи од процес на селекција кандидати од определена етничка заедница или пак да преферира одреден пол, тоа треба да биде посочено како нешто кое со помош на вештачката интелигенција треба да се избегне.

Друг пример е користењето на вештачката интелигенција за мерење на перформанисте на вработените. Доколку компанијата претходно ја мерела успешноста и посветеноста на вработените исклучиво врз основа на бројот на работни часови поминати во канцеларијата во претходниот месец, голема е веројатноста да се дискриминираат вработени кои можеби имале теренски активности кои успешно ги завршиле. Затоа, при дефинирањето на алгоритмите треба да бидат земени предвид други сетови на податоци кои ќе ја дадат проценката без можност за дискриминација.

Кандидатите за вработување од првиот пример и вработените чија успешност се мери во вториот пример имаат право да побараат објаснување за тоа на кој начин била донесена одлуката за нив базирана исклучиво на користење на автоматизирани средства.

Дебатата за тоа кој начин е подобар, посигурен и полесен за контролирање се води долго време помеѓу критичарите на користењето на вештачката интелигенција и оние кои работат на нејзино надградување. Одговорот на критиките генерално е во насока на тоа дека алгоритмите полесно се следат отколку луѓето кои носат одлуки и дека грешките во алгоритмите може прецизно да се идентификуваат и да се отстранат. Повикувањето на потребата за регулација не треба да води кон целосно исклучување на автоматското носење на одлуки туку кон изнаоѓање на алтернативи кои ќе помогнат да се воспостави балансот помеѓу користењето на вештачката интелигенција во процесот и заштитата на човековите права. Една алтернатива која во многу случаи се покажала како функционална е вклучувањето на човечкиот фактор во одреден дел од процесот на носење

на одлука. Ова е тесно поврзано со правото на засегнатото лице за кое се носи одлуката да побара ревидирање на таа одлука со вклученост на човечки фактор.

Секое лице има право да бара да не биде предмет на одлука заснована единствено на автоматска обработка на податоци, особено ако таа одлука предизвикува правни последици за него. Доколку лицето кое било предмет на автоматско носење на одлука побара таа одлука да биде повлечена, тоа мора да биде овозможено. На барање на засегнатото лица, а со помош на човечкиот фактор, одлуката може да биде преиспитана или едноставно лицето да се исклучи од целиот процес.

На овој начин, не само што ќе се испочитува правото на субјектот на личните податоци да не биде предмет на автоматско носење на одлуки, туку ќе биде демонстрирано и почитувањето на транспарентност и правичност на обработката на личните податоци.

Сепак, предизвикот да се пронајде најдобрата комбинација на интеракција помеѓу човекот и вештачката интелигенција при носењето на одлуки останува земајќи ги предвид предностите и маните на секој од нив.

5.3. Проценка на влијанието на заштитата на личните податоци

Проценката на влијанието на заштитата на личните податоци задолжително треба да претходи на развојот на вештачката интелигенција чие користење вклучува и обработка на лични податоци. Агенцијата за заштита на личните податоци има донесено Правилник за процесот на проценка на влијанието на заштитата на личните податоци[33] согласно кој проценката треба да се изврши пред самото започнување на обработката, односно, во процесот на дефинирање на самата алатка која ќе обработува лични податоци со помош на вештачка интелигенција. Освен во процесот на развој, проценката на влијанието на заштитата на личните податоци треба да се врши и дополнително доколку се врши надградување со внесување на сосем нов процес во веќе постоечка алатка која користи вештачка интелигенција.

Со цел давање на јасни насоки за сите оние кои отпочнуваат со нова обработка на лични податоци, Правилникот за проценка на влијанието на заштитата на личните податоци содржи и Листа на видовите на операции за кои се бара да се спроведе проценка[34]. Како што е дефинирано во оваа листа, видовите на операции за кои се бара проценка на влијанието врз заштитата на личните податоци вклучуваат:

- обработка на лични податоци за систематско и сеопфатно профилирање или автоматско донесување одлуки со цел да се извлечат заклучоци и да се донесат одлуки кои произведуваат правно дејство, кои во голема мера влијаат на физичкото лице и/или на повеќе лица или кои помагаат при донесување на одлуки за нечиј пристап до услуга или некој вид на услуга или некоја погодност (на пр., како што се обработка на лични информации во врска со економски или финансиски статус, здравје, лични преференции, интереси, сигурност, однесување, податоци за локација, итн.);
- обработка на посебни категории на лични податоци со цел профилирање или автоматско донесување на одлуки;
- обработка на посебни категории на лични податоци, т.е. податоци што откриваат расно или етничко потекло, политичко мислење, религиозно или филозофско уверување или членство во синдикат, како и обработка на генетски податоци, биометриски податоци со

[33] Правилник за процесот на проценка на влијанието на заштитата на личните податоци. Агенција за заштита на лични податоци.

[34] Листа на видовите операции на обработка за кои се бара проценка на влијанието врз заштитата на личните податоци. Агенција за заштита на лични податоци.

цел единствено идентификување на лицата, здравствени податоци или податоци за сексуалниот живот или сексуалната ориентација на индивидуата;
- обемна обработка на посебни категории на лични податоци или на лични податоци поврзани со казнените осуди и казнените дела;
- обработка на лични податоци на деца со цел профилирање, автоматско одлучување или за цели на маркетинг или за директно понудување на услуги наменети за нив;
- обработка на лични податоци собрани од трети страни (лица), кои се земаат предвид за донесување на одлуки поврзани со склучување, раскинување, одбивање или продолжување на договори за давање на услуги на физички лица;
- обработка на лични податоци со користење на систематско набљудување (мониторинг) на јавно достапен простор во големи размери;
- употреба на нови технологии или технолошки решенија за обработка на лични податоци или со можност за обработка на лични податоци што служат за анализирање или предвидување на економската состојба, здравјето, личните желби или интереси, сигурноста или однесувањето, локацијата или движењето на физичките лица;
- обработка на лични податоци преку поврзување, споредување или вршење на проверка на сличностите од повеќе извори;
- обработка на лични податоци на начин кој што вклучува следење на локацијата или на однесувањето на физичкото лице во случај на систематска обработка на податоците за комуникација (метаподатоци) настанати – генерирани со употреба на телефон, интернет или други средства (канали) за комуникација, како што се GSM, GPS, Wi-Fi, за следење и обработка на податоците за локацијата;
- обработка на лични податоци преку користење на уреди и технологии, кај кои што доколку настане инцидент може да го загрози здравјето на една личност или повеќе лица (субјекти на лични податоци);
- обработка на посебни категории на лични податоци на вработените кои се користат за единствена идентификација на вработените од страна на работодавачот и во други случаи на обработка на податоци за личности – вработени од страна на работодавачот преку користење апликација или систем за следење на нивната работа, движење и комуникација и слично (на пр. обработка на лични податоци за следење на вршењето на работната обврска, движењето, комуникација и сл.).

Со проценката на влијанието на заштитата на личните податоци неопходно е да се идентификува и документира степенот на човечкото влијание во одлучувањето при обработката на личните податоци и во која фаза тоа се случува. Проценката особено треба да вклучи:
- Опис на видовите и целите на обработка на личните податоци, категориите на лични податоци кои ќе се собираат и обработуваат по пат на вештачка интелигенција, роковите на чување, дали ќе има трети страни кои ќе ги користат тие лични податоци и за која цел;
- Анализа на неопходноста и пропорционалноста на обработката која ќе треба да даде одговор на прашањето дали системот за вештачка интелигенција е способен да ги обработува личните податоци за конкретните цели, односно, да покаже дека целта не може да биде постигната доколку не се обработуваат лични податоци;
- Анализа на ризиците за правата и слободите на корисниците што можат да настанат со обработката на нивните лични податоци кои се обработуваат во систем за вештачка интелигенција;
- Предвидени мерки за демонстрирање на усогласеност со прописите за заштита на лични податоци.

Процесот на влијанието на заштитата на личните податоци задолжително треба да биде документиран и да биде поткрепен со план за редовни периодични контроли со цел следење и преиспитување на функционалностите на системот.

5.4. Права на заштита на личните податоци

Право на пристап
На локацијата каде е достапна алатката или апликацијата која користи вештачка интелигенција треба да има објавено информација за приватноста и со тоа да се обезбеди правото на пристап на корисниците. Корисникот има право да добие информации за: целите на обработката, видови на лични податоци кои се обработуваат, корисници на кои личните податоци се даваат на користење, рок на чување, правото на исправка или бришење, ограничување на обработка, приговор, право да поднесе барање до надлежен орган за заштита на личните податоци, постоење на автоматизиран процес на одлучување, вклучувајќи и профилирање.

Право на исправка
Доколку при остварувањето на пристап, корисникот забележи дека неговите лични податоци се неточни или непотполни, има право да побара нивна исправка и надополнување.

Право на бришење
Корисникот има право да побара неговите лични податоци да бидат избришани во случаите кога се исполнети целите заради кои се обработени, ако е повлечена согласноста, а притоа нема друг правен основ, ако приговара на обработката на податоците, ако личните податоци биле незаконски обработени, заради почитување на законска обврска или ако податоците биле собрани во врска со понуда на услуги на информатичко општество.

Право на ограничување на обработката
Корисникот може да бара ограничување (блокирање) на обработката на личните податоци во случаи кога: ја оспорува нивната точност; обработката е незаконита, а тој се противи на бришење на податоците; податоците веќе не се потребни да бидат чувани, но потребни му се на корисникот за остварување на неговите правни барања и кога корисникот вложил приговор, па се чека исходот од приговорот. Кога е ограничена обработката, податоците само се чуваат и не се обработуваат понатаму.

Право на преносливост
Корисникот има право да ги добие своите лични податоци во структуриран, вообичаено користен и машински читлив формат и истите да ги пренесе на друго место, без попречување од страна на системот за вештачка интелигенција. Ова право е применливо кога се врши обработка врз основа на согласност или договор и кога обработката се врши со автоматизирани средства.

Право на приговор
Корисникот има право да поднесе приговор врз основа на конкретна ситуација поврзана со него и тоа кога обработка на личните податоци која се заснова на јавен или легитимен интерес, вклучувајќи и профилирање; обработка на лични податоци се врши за цели на директен маркетинг и профилирање поврзано со директниот маркетинг; обработката на личните податоци се врши за цели на научни или историски истражувања или за статистички цели. Понатамошна обработка на личните податоци е можна, освен ако постојат релевантни легитимни интереси кои преовладуваат над интересите, правата и слободите на корисникот.

Автоматско донесување на поединечни одлуки и профилирање
Корисникот има право да бара да не биде предмет на одлука заснована единствено на автоматска обработка на податоци, врз основа на профилирање, ако таа одлука предизвикува правни последици за него. Ова не се однесува на одлуки кои се засноваат на договор, закон или

согласност. Ако одлуката се заснова на договор или согласност, субјектот има право да бара да биде обезбедена човечка интервенција, да бара да се преиспита одлуката и да изрази личен став.

Право на повлекување на согласноста

На барање на корисникот, обработката на неговите лични податоци може да престане. Повлекувањето на согласноста не влијае на законитоста на обработката која се вршела врз основа на согласност пред отповикувањето. Корисникот треба да има можност да побара повлекување на согласноста на истиот начини, односно, по истиот пат по кој и првично ја дал, па така ако согласноста била дадена по електронски пат, со испраќање на електронска порака или одбирање на соодветно поле, на истиот начин треба да може и да ја повлече.

6. Запознаеност на актерите вклучени во процесот со основните принципи на развој и користење на вештачка интелигенција

Со цел добивање на информации за актуелната состојба во Република Северна Македонија, а како составен дел од ова истражување, беше подготвен прашалник наменет за владините и јавните институции, независните регулатрони тела и агенции, образовните институции и приватните компании кои се вклучени во различни фази од развојот на вештачката интелигенција. Одговорите на прашањата даваат приказ на тоа како вклучените актери работат кон воспоставување на баланс помеѓу брзиот развој на технологијата и носењето на нови и прилагодувањето на постоечките политики, а со цел заштита на човековите права од инвазивното користење на новите технологии со особен фокус на развојот и користењето на вештачката интелигенција.

Прашањата кои беа дел од овој прашалник се:

1. На кој начин Вашата институција/компанија е вклучена во процесот на развој и користење на новите технологии, особено вештачката интелигенција?
2. Проактивноста на институциите/компаниите е клучна за воспоставување на систем и принципи врз кои ќе се планира, развива и понатаму користи вештачката интелигенција. Дали до сега имате направено анализа на потребите за развивање и користење на вештачката интелигенција за Вашата институција/компанија но и потребите на граѓаните чии права ги штитите и кои се сознанијата од таа анализа (подобрување на личните капацитети, начините на остварување на правата, зголемена инклузивност, намалување на економски, социјални, родови нееднаквости, заштита на животна средина, општа добросостојба и слично)?
3. При развојот и користењето на вештачката интелигенција, сите вклучени актери треба да водат особено внимание за почитувањето на законодавството и човековите права низ сите фази на процесот. Ова особено се однесува на почитувањето на слободата на мислата и говорот, достоинството, приватноста и личните податоци, заштитата од дискриминација, социјалната правда и правата на работникот. Дали Вашата институција/компанија учествувала во развој на механизми, политики и насоки за правичен развој и користење на вештачката интелигенција во контекст на заштита на горе наведените права?
4. Транспарентноста и отчетноста за начинот на кој се користи вештачката интелигенција се клучни принципи кои не само што ја отсликуваат отвореноста на институциите туку и им даваат на граѓаните можност да се информираат пред и самите да отпочнат да ја користат вештачката интелигенција. Дали имате воспоставено процедура за информирање на граѓаните за начинот на користење на новите технологии вклучувајќи ја вештачката интелигенција и што точно опфаќа истата (опис на алатката/платформата/апликацијата, собирањето и користењето на лични податоци, начин на остварување на правата во случај на нивно прекшување и слично)?
5. Системите на вештачката интелигенција треба да се стабилни, сигурни и безбедни за користење. Дали Вашата институција/компанија поседува доволно капацитети за да може да направи анализа на ризик пред да отпочне со развивање на одредена алатка која користи вештачка интелигенција и воспостави технички и организациски мерки за заштита од можни безбедносни инциденти?
6. Отчетноста на институциите/компаниите е клучна за правилното функционирање на системите на вештачката интелигенција. Дали можете лесно да ја дефинирате Вашата улога во анализата на потребите, развојот и користењето на вештачката интелигенција?

Прашалникот беше испратен до Кабинетот на Заменик на Претседателот на Владата задолжен за политики за добро владеење, Министерството на информатичко општество и администрација, Фондот за иновации и технолошки развој, Агенцијата за заштита на личните податоци, Факултетот за информатички науки и компјутерско инженерство (ФИНКИ), Факултетот за електротехника и информациски технологии (ФЕИТ), Универзитетот за информатички науки и технологии „Св. Апостол Павле“ (УИНТ), Факултетот за информатика при УГД, СЕМОС Академија, и Brainster.

Со оглед на фактот што одговори беа добиени само од Кабинетот на Заменик на Претседателот на Владата задолжен за политики за добро владеење, Агенцијата за заштита на личните податоци, Универзитетот за информатички науки и технологии „Св. Апостол Павле“ (УИНТ) и СЕМОС Академија, не би можело да се донесат заклучоци за финкционирањето на целиот еко систем за развој на вештачката интелигенција во земјата. Сепак, одговорите даваат одредена слика за перцепцијата и важноста која се дава на развојот на ВИ. Во продолжение се добиените одговори и ставови на институциите.

Кабинет на Заменик на Претседателот на Владата задолжен за политики за добро владеење

Кабинетот на Заменик на Претседателот на Владата задолжен за политики за добро владеење, во рамки на своите надлежности, на развојот и користењето на новите технологии, особено вештачката интелигенција гледа како на алатки за подобрување на ефикасноста, транспарентноста и процесите на донесување одлуки, а во насока на имплементација на принципите на добро владеење. Притоа, за да се обезбедат позитивни резултати треба особено да се внимава на етичките аспекти, доброто планирање и тековното следење.

Кабинетот досега нема направено анализи за потребите за развивање и користење на вештачката интелигенција за сопствени потреби ниту за потребите на граѓаните од причина што надлежностите согласно закон не предвидуваат развој на алатки од било каков тип.

Кабинетот на Заменик на Претседателот на Владата задолжен за политики за добро владеење, согласно своите надлежности, досега не учествувал во развој на механизми, политики и насоки за правичен развој и користење на вештачката интелигенција. Досега, од страна на Кабинетот не е воспоставена процедура која би ги информирала граѓаните за начинот на користње на новите технологии.

Агенција за заштита на личните податоци

Агенцијата за заштита на личните податоци ги следи и почитува „Насоките за вештачка интелигенција и заштитата на личните податоци“, донесени во 2019 година од Комитетот на Конвенцијата 108+. Насоките имаат за цел да им помогнат на креаторите на политики, развивачите на вештачка интелигенција, производителите и давателите на услуги да се осигурат дека апликациите за вештачка интелигенција не го повредуваат правото на заштита на личните податоци.

Исто така, Агенцијата ја следи „Препораката за влијанието на алгоритамските системи врз човековите права“, на Комитетот на министри на Советот на Европа, кој во 2020 година издаде збир на насоки со кои се повикуваат владите да се погрижат да не ги прекршуваат човековите права преку сопствена употреба, развој или набавка на алгоритамски системи. Според препораката, владите, како регулатори, треба да воспостават ефективни и предвидливи законодавни, регулаторни и надзорни рамки кои спречуваат, откриваат, забрануваат и

отстрануваат кршење на човековите права, без разлика дали произлегуваат од јавни или приватни чинители.

Агенцијата до сега нема извршено анализа на потребите за развивање и користење на вештачката интелигенција. Во рамки на ЕУ твининг проектот „Поддршка во имплементација на модернизираната национална рамка за заштита на личните податоци“, Агенцијата има учествувано во изработка на документот „Информататор за Вештачка интелигенција“ како еден вид на алатка чија цел е да ѝ помогне на секоја организација да се подготви за иднината на вештачката интелигенција, да се биде пред технолошкиот развој и подготовка за социо-економските промени што ги носи вештачката интелигенција и обезбедување на соодветна етичка и правна рамка.

На веб страницата на Агенцијата за заштита на личните податоци, граѓаните и контролорите можат да пристапат до информативен дел кое е наменет за објаснување на спроведување на Проценката на влијанието врз заштитата на личните податоци - ПВЛЗП (https://azlp.mk/kontrolori/nasoki/pvzlp/), како клучен процес кој задолжително се спроведува доколку предвидената обработка е веројатно дека ќе создаде висок ризик за правата и слободите на физичките лица, како што е случајот со развивање на системи за вештачка интелигенција. Информативниот сегмент објаснува:

1. Случаи во кои е задолжителна ПВЗЛП
2. Случаи во кои не е задолжителна ПВЗЛП
3. Улоги и одговорности
4. Барање мислење од субјектот на личните податоци
5. Основни карактеристики на ПВЗЛП
6. Методологија за спроведување на ПВЗЛП
7. Фази на спроведување на ПВЗЛП
8. Што да направите ако по извршената ПВЗП ризикот е сè уште неприфатлив
9. Дали ПВЗП треба редовно да се преиспитува?
10. Објавување на ПВЗЛП

На веб страницата има и дел кој се однесува на Правата на физичките лица во однос на заштитата на личните податоци (https://azlp.mk/vashite-prava/), тука вклучувајќи го: Правото да не се биде предмет на одлука која се заснова единствено на автоматизирана обработка, вклучувајќи и профилирање, кое е особено важно бидејќи по својата природа, алатките за вештачка интелигенција често вклучуваат автоматизирани одлуки кои се носат врз основа на лични податоци.

Агенцијата има значителен недостаток од човечки ресурси, односно, на обучен и стручен кадар кој располага со знаења од областа на заштитата на личните податоци и посебно експерти од областа на информатичките науки кои се клучни за имплементација на легислативата за заштита на личните податоци. Останува потребата да се обезбедат нови вработувања, како и да се обезбедат дополнителни финансиски средства за мотивирање, задржување и понатамошно професионално усовршување на човечките ресурси во Агенцијата со цел да се спречи заминување на постојниот стручен и обучен кадар.

Вештачката интелигенција често се користи за автоматско донесување на поединечни одлуки, вклучувајќи и профилирање, без обезбедување на човечка интервенција. Агенцијата, според член 26 од Законот за заштита на личните податоци, го регулира правото на физичкото лице да не биде предмет на одлука заснована единствено на автоматизирана обработка, вклучувајќи го и профилирањето што предизвикува правни последици за него или на сличен начин значително влијае на него. На организациите им е дозволено автоматски да донесат поединечни одлуки,

вклучувајќи го и профилирање само во следните случаи: одлуката е потребна за склучување или извршување на договор помеѓу организацијата и физичкото лице, одлуката е дозволена со закон (на пример, за целите на измама или даночно затајување), или врз основа на изречна согласност на физичкото лице.

Развојот на системи за вештачка интелигенција, во некои случаи, бара реализација на Проценка на влијанието на заштитата на личните податоци. Според член 39 од Законот за заштита на личните податоци, при користење на нови технологии за некој вид на обработка, која според природата, обемот, контекстот и целите на обработката, постои веројатност истата да предизвика висок ризик за правата и слободите на физичките лица пред да биде извршена обработката, контролорот е должен да изврши проценка на влијанието на предвидените операции на обработката во однос на заштитата на личните податоци. Агенцијата ги има донесено и овие подзаконски акти: „Листа на видовите операции на обработка за кои се бара проценка на влијанието врз заштитата на личните податоци“; и „Листа на видовите операции на обработка за кои не се бара проценка на влијанието врз заштитата на личните податоци“ кои се исто така објавени на официјалната веб страница http://www.azlp.mk/.

Универзитет за информатички науки и технологии „Св. Апостол Павле“ (УИНТ) – Охрид

Универзитетот за информатички науки и технологии „Св. Апостол Павле“ (УИНТ) – Охрид, како високообразовна институција специјализирана во областа на информатичките технологии е директно вклучена во примена и развој на новите технологии, вклучително академскиот пристап на користење на вештачката интелигенција. Најпрво во делот на настава, предметот вештачката интелигенција се учи на додипломски студии во 3 година на факултетите на УИНТ. Во рамките на овој предмет студентите се стекнуваат со основни вештини за носење рационални одлуки кај интелигентните агенти, како и примена на основите алгоритми за движење низ просторот на состојби. Дополнително, во текот на своето образование студентите се оспособуваат да ги применат во моделите за машинско учење кои интелигентните агенти ги користат за одлучување. Во рамки на научно-истражувачките активности на академските студии за втор циклус на студии (магистерски) се изучуваат напредни алгоритми на машинско учење кои се применуваат решавање на комплексни проблеми од секојдневието. Во рамки на наставно-научните активности се вклучени и гостувања на докажани професионалци од пракса и истражувачи кои ја користат вештачката интелигенција за развој на модели за рана детекција и предвидување на решенија на одредени проблеми од секојдневието.

Студентите и академскиот кадар во рамки на УИНТ активно се вклучени во развој на дигитални решенија кои ги користат моделите на вештачката интелигенција. На тој начини УИНТ активно придонесува во развој на апликации кои се базирани на вештачката интелигенција. Дополнително, во рамки на научно-истражувачките активности (дипломски, магистерски и стручни трудови) секојдневно УИНТ придонесуваат за развој на нови иновативни алгоритми за машинско учење.

УИНТ е активно вклучен во развојот на вештачката интелигенција, преку апликативни решенија и нови иновативни алгоритми во рамки на научно-истражувачките активности. Како академска институција, УИНТ се залага да се следат етичките принципи за примена на вештачката интелигенција. Во таа насока, се внимава примената на вештачката интелигенција да има пред сè научен придонес и подобрување на секојдневните активности на луѓето, од аспект на инклузивност, економски и подобри социјални решенија. УИНТ нема направено анализа на потребите за развивање и користење на вештачката интелигенција во рамки на универзитетот.

Академската заедница на УИНТ секогаш ги зема предвид и се консултира со сите засегнати чинители во процесот на развој на апликативни решенија или при споредување на научно-истражувачка дејност. Во таа насока, се почитува етичкиот кодекс на академската институција, а дополнително сите засегнати учесници во развојот и истражувањето се соодветно консултирани. При обработката на податоците (датасетови) се внимава да не содржат лични податоци, а по потреба се врши анонимизација или псевдоанонимизација на податоците. УИНТ до сега не бил консултиран во развој на механизми, политики и насоки за правичен развој и користење на вештачката интелигенција.

Академската заедница на УИНТ тежнее кон примена на софтвер со отворен код, каде што е од примарно значење отвореноста и транспарентноста во развојот на апликативни решенија и нови иновативни алгоритми во рамки на научно-истражувачките активности. Задолжителен дел при развој на софтверските системи е собирање на корисничките барања, во таа насока се прави комплетна спецификација на дигиталното решение кое се изработува на УИНТ. Крајното апликативно решение покрај примената на новите технологии, задолжително е проследено со корисничко упатство во кое точно е наведено како се користи апликацијата. Особен акцент се става на процесот на регистрација и автентификација на корисниците на системот, со посебно предупредување како да се чуваат личните податоци на корисниците. Од безбедносен аспект, препораката е секогаш да се енкриптираат лозинките и личните податоци во базата на податоци. Дополнително, при користење на облак-базираните системи, анг. cloud computing, се внимава каде се чуваат податоците и кои закони важат во соодветната земја каде е сместен податочниот центар. Секако препорака е дека треба да се следат законските регулативи за обработката на податоците во согласност со соодветните национални регулативи.

Академската заедница на УИНТ се стреми за сите апликативни решенија да има воспоставена процедура за анализа на ризици. Конкретно до сега нема воспоставено правилник/процедура за задолжителна анализа на ризици при развој на апликативни решенија. УИНТ има доволно знаење и капацитети да се воспостават технички процедури за заштита и справување со ризици и можни безбедносни инциденти на системите кои користат вештачка интелигенција.

Како јавна високообразовна институција, УИНТ е законски обврзана да биде целосно транспарентна во своето работење. Земајќи предвид дека УИНТ е академска институција која е специјализирана за информатички науки и технологии, од значење е да биде вклучена и консултирана при носењето на законски предлози, правилници и организациони мерки на национално ниво. Доколку се формира национално тело или работна група која ќе е задолжена за регулација на развојот на вештачката интелигенција, претставници од УИНТ треба да се директно вклучени при изработка на нацрт решенија за регулација на развојот и користењето на вештачката интелигенција. УИНТ има свој претставници во МАРНЕТ, Македонска академска истражувачка мрежа, која е обезбедување на услуги на национално ниво и меѓународно поврзување на Македонската академска истражувачка и образовна заедница и поддршка на нивните истражувачки и образовни активности; промовирање и дисеминирање на употребата на информациските и комуникациските технологии посебно во академскиот и истражувачкиот сектор; одржување и управување со националниот ДНС, меѓународно претставување и членство; водење на политика и развој на националната академска мрежа.

<u>CEMOC Едукација</u>

Мисијата на CEMOC Едукација е своите курикулими секогаш да ги гради во согласност со развојот на информатичката технологија и во партнерство со најголемите светски компании со развиени курикулуми за учење. Проширувањето на областите на едукација, со динамика

диктирана од развојот на општеството и пазарот и развивањето на нови сервиси и услуги се секогаш приоритет.

Академија за вештачка интелигенција на CEMOC Едукација е дизајнирана така што го гради знаењето почнувајќи од основата па сè до нивото на сертифициран Artificial Intelligence Practitioner (применувач на вештачка интелигенција).

Проактивноста на компаниите е клучна за воспоставување на систем и принципи врз кои ќе се планира, развива и понатаму користи вештачката инелигенција. Заклучоците до кои има дојдено тимот на CEMOC Едукација од спроведените анализи за потребите за развој на алатки и системи кои користат вештачка интелигенција упатуваат на потребата за подобрени лични капацитети, а со тоа и подобрување на капацитетите на компаниите вклучувајќи ги техничките решенија каде е имплементирана вештачката интелигенција во секојдневната работа. Потребите се различни, од користење на low-code алатки за поголема ефективност, до имплементација на решенија базирани на вештачка интелигенција во процесот на работа.

Постоечките правила, механизми, политики и насоки за правичен развој и користење на вештачката интелигенција во контекст на заштита на човековите права се составен дел од наставните содржини. CEMOC Едукација работи со светски вендори како партнери и спроведува сертифицирана програма низ наставните содржини. Поаѓајќи од подигање на свеста за користење на вештачката интелигенција па се до задолжителните компаниски рамки за почитување на основни начела кои секој вендор ги применува во развој на своите решенија се дел од содржините на обуките на CEMOC Едукација.

CEMOC Едукација нема посебна процедура за информирање на корисниците за начинот на користење на новите технологии вклучувајќи ја вештачката интелигенција и што точно опфаќа истата но има поставено веб страница која обработува содржини исклучиво на тема вештачка интелигенција.

CEMOC Едукација ја препознава важноста на градењето на стабилни, сигурни и безбедни за користење системи на вештачка интелигенција и креирање на документација за секој производ/услуга која користи вештачка интелигенција со детална спецификација за евентуални ризици за правата на корисниците.

Академијата за вештачка интелигенција се развива во насока на прогрес, односно, постојано следење на новите достигнувања во областа и прилагодување на содржините во насока на максимална подготовка на слушателите за материјата. Воведувањето на модули кои ќе ги следат сите измени и новини во областа е од клучно значење за конкуретна предност на слушателите како професионалци кои практицираат вештачка интелигенција.

7. Етички аспекти на вештачката интелигенција

Вештачката интелигенција, покрај огромните можности за иновации во различни сектори, истовремено отвора и различни прашања за заштитата на основните човекови права и етичките стандарди и постулати. Постигнувањето на вистинскиот баланс помеѓу иновацијата и одговорноста е од суштинско значење за обезбедување на дигитално опкружување кој вклучува решавање на проблеми како што се приватноста, пристрасноста, дискриминацијата и етичката употреба на вештачката интелигенција за заштита на човековите права. Како што вештачката интелигенција продолжува да навлегува во нашето секојдневно живеење, потребата за заштита на човековите права без задушување на технолошкиот напредок станува најголем предизвик.

Покрај позитивното право, големо значење за заштитата на човековите права имаат и треба да имаат универзалните деловни, етичките и моралните правила и стандарди. Информатичката револуција предизвика драстични промени во општеството и, соодветно на тоа, правото мора да се промени и приспособи, за да ги вклучи сите нови облици на однесување на правните субјекти и да ги заштити највисоките вредности во општеството, меѓу кои се и човековите права. Бидејќи технолошките промени се исклучително динамични, а позитивното право не може толку динамично да се менува, улогата на универзалните деловни, етичките и моралните правила и стандарди усвоени од деловните и професионалните здруженија, како и од многу други национални и меѓународни организации, станува сè поважна. Овој тип на правила на однесување овие организации можат да ги донесат со соодветна динамика и со почитување на највисоките професионални стандарди, како и стандардите од областа на човековите права. Кодексот на правила наменет за инженерите кои креираат софтвер од областа на вештачката интелигенција треба да ги содржи највисоките етички и морални стандарди, со цел да се спречи кршењето на човековите права што е можно повеќе. Ваквите кодекси претставуваат вреден ресурс и индикатор за идни промени во позитивните законски прописи.

Серија резолуции поврзани со вештачката интелигенција беа усвоени од Европскиот парламент во 2020 година: Резолуција за етичките аспекти на вештачката интелигенција, роботиката и сродните технологии[35], Резолуцијата за режимот на граѓанска одговорност за вештачката интелигенција[36], Резолуцијата за правата на интелектуална сопственост поврзани со развојот на вештачка интелигенција[37]. UNESCO во ноември 2021 година усвои Препораки за етика кај вештачката интелигенција. Како основни принципи во Препораките се наведени: пропорционалност и непричинување штета, безбедност и сигурност, правичност и недискриминација, одржливост, право на приватност и заштита на податоците, човечки надзор и посветеност, транспарентност и објаснивост, одговорност, свесност и писменост, и адаптивно управување и соработка со повеќе засегнати страни. Како што претходно веќе беше спомнато, Европската комисија во 2023 година достави до Европската Унија предлог за регулаторна рамка за вештачка интелигенција. Претходно, комесарот за човекови права при Советот на Европа издаде препораки од десет точки за вештачката интелигенција и човековите права, кои се надоврзуваат на она што Советот на Европа веќе го направи во оваа област, особено преку Европската етичка повелба за употреба на вештачката интелигенција во правосудството, Декларацијата на Советот на министри за манипулативните можности на алгоритамските процеси, Студијата за димензиите на човековите права во техниките за автоматска обработка на податоци и можните регулаторни импликации, како и извештајот на специјалниот известувач

[35] Усвоени текстови - Рамка на етички аспекти на вештачката интелигенција, роботиката и сродните технологии. Европски парламент., преглед на 22.12.2023 год.

[36] Усвоени текстови - Режим на граѓанска одговорност за вештачка интелигенција. Европски парламент., преглед на 22.12.2023 год.

[37] Усвоени текстови - adopted - Права на интелектуална сопственост за развој на вештачка интелигенција. Европски парламент., преглед на 22.12.2023 год.

на Обединетите нации за промоцијата и заштита на слободата на мислата и изразувањето, кој ги разгледува импликациите на технологиите за вештачка интелигенција врз човековите права во информатичкото општество. Исто така, во согласност со иницијативата за безбедно развивање на вештачка интелигенција за човечката раса, голем број граѓански организации и експерти за вештачка интелигенција креираа низа одлуки, препораки и декларации во кои се даваат насоки за безбеден развој на вештачката интелигенција. Со оглед на новите предизвици, во смисла дека нејзиниот развој „би можел“ да ја надмине човечката интелигенција, се препорачува да се додаде етичка компонента на комплицираните одлуки преку човечкиот „final touch“.

Без да се намали важноста на другите сфаќања и начела, издвоени се следните **етички принципи** кои се препознаваат како појдовна точка за создавање, примена и употреба на системи за вештачка интелигенција, а кои треба да претставуваат применливи, достојни и неприкосновени правила за зачувување на човечкото достоинство поради нивната сигурност, доверливост и одговорност кон луѓето:

- **Објаснивост и проверливост** - Една од основните карактеристики на човековата свест е да ја согледува околината, да бара одговори на прашања и објаснувања зошто и како нешто е или не е. Оваа карактеристика влијаела на еволуцијата на човекот и развојот на науката, а со тоа и на вештачката интелигенција. Потребата на човекот да ги разбере и да ги разјасни работите ја најде својата основа во овој принцип. Разбирливоста (јасноста, објаснивоста) значи дека сите процеси: развој, тестирање, пуштање во употреба следење на системот и исклучување мора да бидат транспарентни и да можат лесно да се објаснат. Целта и можностите на самиот систем на вештачка интелигенција мора да се објаснат, а особено одлуките (препораките) што ги носи системот. Доколку одредени резултати од работата на системот на вештачката интелигенција не можат да се објаснат и да се разберат, потребно е да се означат како систем со модел на „црна кутија[38]“. Проверливоста е комплементарен елемент на овој принцип, кој осигурува дека системот може да се провери во сите процеси, односно во текот на целиот свој животен циклус. Проверливоста ги вклучува активностите и процедурите за проверка на системите за вештачка интелигенција за време на тестирањето и имплементацијата, како и проверка на краткорочното и долгорочното влијание што таквиот систем го има врз корисниците.

- **Достоинство** - Обврска на сите членови на општеството е меѓусебно да го почитуваат и да го промовираат и штитат ова право како едно од основните и неприкосновени права на секое човечко суштество. Секој поединец има право да го заштити сопственото достоинство, а повредата или непочитувањето на ова право се санкционира со закон. Човечкото достоинство треба да се сфати како фундаментален принцип и основно човеково право кој се фокусира на зачувување на човековиот интегритет. Врз основа на таа премиса, корисниците на системите на вештачката интелигенција треба во секое време, без оглед на фазата во која се наоѓа конкретното решение за вештачка интелигенција (развој, примена или употреба), да ја имаат предвид личноста на човекот и неговиот интегритет како централен концепт и императив. Во таа насока, неопходно е да се развиваат системи кои во секоја фаза задолжително мора да ја почитуваат личноста на човекот, неговата слобода и автономија. Почитувањето на човечката личност значи создавање систем кој ќе ги почитува когнитивните, социјалните и културните карактеристики на секој поединец. Системите за вештачка интелигенција мора да бидат во согласност со горенаведеното и потребно е да се води сметка на кој било начин да не доведат до потчинување на човекот на функциите на системот, како и до загрозување на неговото достоинство и интегритет. За да се обезбеди почитување на принципот на

[38] Црна кутија. Tech Target., преглед на 18.12.2023 год.

достоинство, системите за вештачка интелигенција не смеат во процесите на работа и примена грубо да ја игнорираат или да ја занемаруваат автономијата на човечкиот избор.

- **Забрана за предизвикување штета** - Системот за вештачка интелигенција мора да ги почитува безбедносните стандарди, односно да содржи соодветни механизми кои ќе спречат предизвикување штета на луѓето и нивниот имот. Во случај да настане штета, таа мора да се санира во најкус можен рок, а на оштетеното лице да му се надомести на начин утврден со закон. Законот за облигационите односи го уредува концептот на штета како „намалување на нечиј имот (обична штета) и спречување на негово зголемување (загубена корист), како и нанесување физичка или психичка болка или страв на друг (нематеријална штета)[39], и утврдува дека секое лице е должно да се воздржува од дејствија кои можат да предизвикаат штета на други.” Покрај граѓанската одговорност, законот ја препознава и кривичната и прекршочната одговорност и на физичките и на правните лица за штетата што ја предизвикуваат на друго лице. Кривичниот законик предвидува голем број кривични дела, од кои важно е да се споменат кривичните дела против животот и телото[40], имотот на луѓето[41], против слободите и правата на човекот и граѓанинот[42]. Посебно внимание треба да се посвети на заштитата на чувствителните категории како што се старите лица, лицата со попреченост, децата, бремените жени итн., како и категориите кои се во понеповолна положба (на пример: работник - работодавач, потрошувач - трговец итн.). Системите за вештачка интелигенција мора да се користат на безбеден начин, односно да бидат сигурни и безбедни, а нивната употреба за нечесни и злонамерни цели треба да се спречи.

- **Правичност** - Овој принцип се однесува на заштитата на правата и интегритетот од дискриминација, особено врз чувствителните категории (на пример, лицата со попреченост). Самиот поим, поради неговата сестраност, има различни толкувања во бројни сфери на општественото живеење. На пример, во здравствената заштита, принципот на правичност подразбира забрана на дискриминација во обезбедувањето здравствена заштита врз основа на раса, пол, сексуална ориентација и родов идентитет, возраст, националност, социјално потекло, религија, политичко или друго убедување, имотна состојба, култура, јазик, здравствена состојба, вид на болест, ментална или физичка попреченост, како и други лични карактеристики кои можат да бидат причина за дискриминација и стигматизација. Исто така, системите за вештачка интелигенција мора да спречат дискриминација во процесот на нивното користење. Системите за вештачка интелигенција треба да овозможат еднакви можности за сите лица, како во однос на пристапот до образование, стоки и услуги и технологии, така и во превенцијата, односно да спречат измама од страна на лица кои користат системи за вештачка интелигенција при донесување одлуки базирани на препораките на системот за вештачка интелигенција. Правичната употреба на системите за вештачка интелигенција може да доведе до зголемување на правичноста во општеството како целина, како и до намалување на разликите што постојат меѓу поединците во однос на социјалниот, економскиот и образовниот статус.

[39] Член 142, Закон за облигационите односи
[40] Глава 15, Кривичен законик
[41] Глава 26, Кривичен законик
[42] Глава 16, Кривичен законик

- **Етичка и општествено одговорна употреба на податоците** – Овој принцип ја подразбира обврската на ракувачите и обработувачите на податоци, особено кога информациите се користат за предвидливи цели во процесите на донесување одлуки, да го земат предвид веројатното влијание на планираната обработка на големи податоци врз фундаменталните права и слободи и особено на обврската за почитување на заштитата на личните податоци. Обработката на податоците не треба да биде во конфликт со етичките вредности, така што ракувачите со податоци би можеле да формираат етичка комисија и да воспостават етички кодекс, со цел да ги идентификуваат етичките вредности што треба да се заштитат при обработката на податоците. Всушност, системот за вештачка интелигенција може да биде етички едуциран како што ние ќе го програмираме. Се препорачува развивачите на вештачка интелигенција, производителите на системи за вештачка интелигенција и давателите на услуги за вештачка интелигенција да соработуваат со независни експерти и независни академски експерти од различни области, со цел да се заштитат човековите права и етичките принципи при дизајнирање системи за вештачка интелигенција, особено во областа на правдата и предвидување и откривање на кривични дела. Насоките нагласуваат дека сите производи и услуги за вештачка интелигенција треба да бидат дизајнирани на таков начин што ќе го осигураат правото на поединците да не бидат предмет на одлуки донесени со автоматизирани средства, без да се земат предвид нивните ставови. Сите апликации за вештачка интелигенција треба да обезбедат усогласеност со принципите за заштита на човековите права и заштита на податоците во текот на целиот животен циклус на тие апликации. Корисниците секогаш треба да бидат информирани дека имаат интеракција со апликации за вештачка интелигенција и имаат право да добиваат информации за основните принципи на обработка на податоците што важат за нив. Корисниците на апликации за вештачка интелигенција мора да имаат право на приговор во врска со обработката на податоците што влијае на ставот и личниот развој на поединецот.

Со подобрување на предвидувањето, оптимизирање на операциите и распределбата на ресурсите и персонализирање на дигиталните решенија достапни за поединци и организации, употребата на вештачка интелигенција може да обезбеди клучни конкурентски предности за компаниите и да ги поддржи социјално и еколошки корисни резултати во здравството, земјоделството, безбедноста на храната, образование и обука, медиуми, спорт, култура, управување со инфраструктура, енергија, транспорт и логистика, управување со кризи, јавни услуги, безбедност, правда, ресурси и енергетска ефикасност, мониторинг на животната средина, зачувување и обновување на биодиверзитетот и екосистемите и ублажување на климатските промени и адаптација.

Како позитивен пример од нашето блиско соседство, кој со мали прилагодувања можеме да го преземеме и примениме во Република Северна Македонија, можеме да ги посочиме Етичките насоки за развој, примена и употреба на доверлива и одговорна вештачка интелигенција[43] кои беа усвоени од страна на Владата на Република Србија во март 2023 година, а со цел воспоставување на национален превентивен механизам кој ќе овозможи одговорен развој на вештачката интелигенција и начините на верификација на системите базирани на машинско учење согласно највисоките етички и безбедносни стандарди. Целта на овие Етички насоки е да се превенира загрозување или маргинализирање на граѓаните и нивните активности од страна на системите и процесите на вештачка интелигенција и да се спречи нарушување на слободата на мислење, дејствување или одлучување на поединците, до таа мерка каде што правото и етичките стандарди кои ги чуваат и гарантираат таквите права можат да бидат намалени или

[43] Нацрт – Етички упатства за развој, примена и употреба доверлива и одговорна ВИ. Национална платформа за вештачка интелигенција., преглед на 18.12.2023 год.

целосно заборавени. Како приоритети и највисоки вредности во самите Етички насоки се претставени заштитата на личните податоци, заштитата од дискриминација при употребата на машинското учење и воспоставување на одговорен развој на вештачката интелигенција согласно меѓународните етички принципи и обврзувачките и упатувачките меѓународни правни акти.

Како резултат на глобалните трендови во дигитализацијата, на иницијатива на Фондот за иноваци и технолошки развој, беше отпочнат процесот за креирање на Националната стратегија за вештачка интелигенција во Република Северна Македонија со цел да се даде можност на домашните, иновативни стартап компании да ги реализираат своите идеи и проекти, соодветно образование како и пристап до современа опрема. За креирање на Националната стратегија за вештачка интелигенција, во септември 2021 година, беше формирана работна група[44] која вклучува домашни експерти но и успешни, глобално признаени професионалци од оваа област, а кои работат во светски познати компании и универзитети. Фондацијата за интернет и општество – Метаморфозис е вклучена во оваа работна група, а преку определбата за континуирано реализирање на активности за истражување, обука, промоција и застапување за етичкото користење на дигиталните технологии и заштитата на човековите права на интернет ќе даде свој придонес во работата на оваа група во следниот период. Во иста насока, Метаморфозис го започнува процесот за формирање Коалиција за одговорна вештачка интелигенција[45] која треба да собере чинители од повеќе сектори. Коалицијата треба да вклучи не само граѓански организации, туку и експерти од академската сфера и од бизнис секторот за да се обезбеди инклузивен и сеопфатен пристап во оформувањето на нашата иднина, бидејќи токму вештачката интелигенција е таа што ја обликува нашата иднина.

Етичките стандарди за развој, имплементација и употреба на доверлива и одговорна вештачка интелигенција се подготвуваат со намера да се обезбеди рамка и да бидат дадени насоки за работата на сите учесници во екосистемот на вештачката интелигенција. Во очекување на посилна правна рамка, која допрва го добива својот облик во Европската Унија, овие стандарди, принципи и насоки овозможуваат понатамошен развој во оваа област. Вештачката интелигенција треба да се користи за доброто на целата заедница. Системите за вештачка интелигенција треба да служат за одржување и негување на демократските процеси и почитување на плуралноста на вредностите и животниот избор на поединците. Етичките стандарди даваат основа за поширока примена на вештачката интелигенција во одлучувањето и во обликувањето на општествените промени, зголемувањето на знаењето и натамошниот економски напредок на општеството како целина. На крајот, со оглед на обемот и сложеноста на глобалната политика и економија, како и нашето знаење за човечката природа, би било крајно наивно да се потпреме едноставно на спонтано и доброволно етичко однесување на поединци и корпорации за да се обезбеди правичност или да се зајакне човечкото достоинство. Постоечките регулативи, и сите други решенија што ќе се донесат, во комбинација со нивната одговорна примена, се неопходни за да го насочат и регулираат нашето однесување и да обезбедат владеење на правото и почитување на законите.

[44] Национална AI стратегија. Фонд за иновации и технолошки развој., преглед 27.12.2023 год.
[45] Фондацијата Метаморфозис создава коалиција за одговорна вештачка интелигенција. Новинска агенција Мета.мк., преглед 27.12.2023 год.

Користена литература

1. Вештачката интелигенција – застрашувачки корисна, но и опасна алатка во човечките раце. Новинска агенција Мета.мк.
2. Општа регулатива за заштита на податоците (GDPR)
3. Компјутерски машини и интелигенција. А. М. Тјуринг.
4. Туринг тест. Станфорд енциклопедија на филозофија.
5. Логички теоретичар објасни - Сè што треба да знаете. History Computer.
6. Перцепции – Вовед во компјутерска геометрија | MIT Press, М. Мински, С. А. Пејперт.
7. Deep Blue. Chess.
8. Ватсон, шампион на Jeopardy! IBM.
9. Вештачката интелигенција на суперкомпјутерот Baidu Minwa ги надминува Google, Microsoft и луѓето во препознавањето на слики. Ентони Катбертсон.
10. Вештачка интелигенција: AlphaGo на Гугл го победи Go мастерот Lee Se-dol. BBS News.
11. Language Learning Models.
12. Вештачка интелигенција (ВИ). Стејт департментот на САД.
13. Вештачка интелигенција. Британика.
14. Што е вештачка интелигенција (ВИ)? IBM.
15. ЕУ ВИ Акт: првата регулатива за вештачка интелигенција,. Европски парламент.
16. Преглед на ВИ принципи. ОЕЦД.
17. Европратениците се подготвени да преговараат за први правила за безбедна и транспарентна вештачка интелигенција. Европски парламент.
18. Акт на ЕУ за вештачка интелигенција: прва регулатива на оваа област. Европски парламент.
19. Вербативен извештај на постапката - Закон за вештачка интелигенција (A9-0188/2023). Европски парламент.
20. Водич за човекови права за корисници на Интернет. Совет на Европа.
21. Европски дигитални права и принципи. Европска комисија.
22. Закон за заштита на личните податоци, Службен весник на РСМ бр.42/20 и бр.294/21
23. Агенција за заштита на лични податоци
24. Методологија за проценка на влијанието на заштитата на личните податоци. Фондација Метаморфозис.
25. АДА - Првиот дигитален асистент во Владата на Република Северна Македонија
26. Универзална декларација за човекови права - англиски. Обединети нации.
27. Директива на Советот 2000/78/ЕЗ од 27 ноември 2000 година за воспоставување општа рамка за еднаков третман во вработувањето и професијата. EUR-Lex.
28. Влијанието на вештачката интелигенција врз Општата регулатива за заштита на податоците. Единица за научно предвидување на службата за истражување на Европскиот парламент.
29. Правилник за процесот на проценка на влијанието на заштитата на личните податоци. Агенција за заштита на лични податоци.
30. Листа на видовите операции на обработка за кои се бара проценка на влијанието врз заштитата на личните податоци. Агенција за заштита на лични податоци.
31. Усвоени текстови - Рамка на етички аспекти на вештачката интелигенција, роботиката и сродните технологии. Европски парламент.
32. Усвоени текстови - Режим на граѓанска одговорност за вештачка интелигенција. Европски парламент.
33. Усвоени текстови - adopted - Права на интелектуална сопственост за развој на вештачка интелигенција. Европски парламент.
34. Црна кутија. Tech Target.
35. Нацрт – Етички упатства за развој, примена и употреба доверлива и одговорна ВИ. Национална платформа за вештачка интелигенција.
36. Национална AI стратегија. Фонд за иновации и технолошки развој.

37. Фондацијата Метаморфозис создава коалиција за одговорна вештачка интелигенција. Новинска агенција Мета.мк.
38. Алан Тјуринг, „*Computing Machinery and Intelligence*“
39. Џон Мекарти, конференција за вештачка интелигенција на колеџот Дартмут.
40. МарвинИнски и Сејмур Пеперт, „Perceptrons“
41. Закон за облигационите односи, Службен весник на РСМ бр. 154/2023
42. Кривичен законик, Службен весник на РСМ бр. 08-4337/1