

ВПЛИВ ГРОМАДСЬКОЇ ДУМКИ НА ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ У ПРИЙНЯТТІ ПОЛІТИЧНИХ РІШЕНЬ

Дедалі ширше використання штучного інтелекту та алгоритмів при прийнятті рішень ставить низку питань. Насамперед йдеться про алгоритмічну упередженість та непрозорість процесу. На прикладі референдуму у Каліфорнії в 2020 році щодо запровадження алгоритмічної оцінки при прийнятті рішення про звільнення з-під арешту під заставу.

Виділено та проаналізовано три основних чинника, що визначають важливість позитивного ставлення громадян до використання інструментів ШІ у прийнятті політичних рішень: 1) визначення майбутнього ШІ; 2) надмірна політизація питання та бойкот; 3) легітимність прийнятих рішень.

Розглянуто три ключових сфери, де використання ШІ найбільше вплине на ставлення громадськості, зокрема розпізнавання облич, алгоритми персоналізації та автоматизація робочих місць. На основі аналізу результатів опитувань громадської думки визначено чинники, що впливають на ставлення громадян до ШІ.

Ключові слова: штучний інтелект, легітимність, демократія, розпізнавання облич, алгоритмічна упередженість, автоматизація.

Vitushko Maksym. The influence of public opinion on the use of artificial intelligence in political decision-making

The increasing use of artificial intelligence and algorithms in decision-making raises a number of issues. First and foremost, we are talking about algorithmic bias and lack of transparency in the process. The author uses the example of the 2020 referendum in California on the introduction of algorithmic assessment when deciding whether to release a person from bail.

The author identifies and analyses three main factors that determine the importance of a positive attitude of citizens towards the use of AI tools in political decision-making: 1) determination of the future of AI; 2) excessive politicalisation of the issue and boycott; 3) legitimacy of the decisions made.

Three key areas where the use of AI will have the greatest impact on public attitudes are considered, including face recognition, personalisation algorithms,

and workplace automation. Based on the analysis of the results of public opinion polls, the factors influencing the attitude of citizens towards AI are identified.

Key words: *artificial intelligence, legitimacy, democracy, face recognition, algorithmic bias, automation.*

Стурбованість щодо ефективності, упередженості та справедливості використання тучного інтелекту (ШІ) та алгоритмічного прийняття рішень не є чимсь новим, але вона посилилася, оскільки алгоритмічне прийняття рішень дедалі більше поширюється в державній політиці та інших контекстах прийняття рішень з високими ставками, таких як кримінальне правосуддя [1] та система охорони здоров'я [2].

Останнім часом з'явилася низка публікацій, що висвітлюють недоліки в роботі алгоритмічних систем прийняття рішень, насамперед тих, що пов'язані з несправедливістю та/або упередженістю щодо певних груп. У 2016 році був опублікований аналіз алгоритму, що використовувався для генерації балів ризику для підсудних у кримінальних справах. За його підсумками було зроблено висновок, що алгоритм був упередженим щодо чорношкірих підсудних [3]. У 2019 році була опублікована стаття [4], яка навела докази расової упередженості алгоритму, що використовується медичною системою США для розподілу ресурсів спеціалізованої медичної допомоги між пацієнтами.

Показовою стала ситуація щодо впровадження елементів штучного інтелекту при прийнятті рішень в одному з центрів технологічного розвитку США, де знаходиться Кремнієва долина – у Каліфорнії. У 2020 році на референдум серед низки пропозицій була винесена ініціатива про заміну грошової застави на алгоритмічну оцінку ризиків. Згідно з Конституцією штату, особи, заарештовані та поміщені до окружної в'язниці, за винятком певних тяжких злочинів, мають право на звільнення до суду [5]. Конституція визначає, що ці люди повинні бути звільнені за умов, які не є надмірними. Приймаючи рішення про звільнення особи до суду, суди першої інстанції повинні враховувати: тяжкість злочину, в якому обвинувачується особа; наявність у неї попередніх судимостей і ймовірність того, що вона з'явиться до суду. При цьому суди можуть використовувати різну інформацію, включаючи оцінку ризиків.

Щоб допомогти у прийнятті рішень про те, чи звільняти людей до суду, більшість судів і округів використовують інструменти для оцінки ризику (або ймовірності) того, що звільнена особа вчинить новий злочин або не з'явиться до суду. Ці інструменти були розроблені на основі досліджень, які показують, що люди з певними рисами мають більше шансів скоїти новий злочин або не з'явитися до суду. На основі певних характеристик людини їй нараховуються бали. Наприклад, більше балів отримують особи, яким не виповнилось 22 роки, оскільки вони більш схильні до скоєння злочинів, ніж люди старшого віку. Аналогічно, ті, хто кілька разів не з'являвся до суду в минулому, з меншою ймовірністю з'являться в майбутньому й отримують більше балів. Рівень ризику особи визначається загальною кількістю отриманих балів, який згодом використовується для прийняття рішення про те, чи слід звільнити особу і за яких умов.

На голосування виносилася пропозиція про автоматичне звільнення протягом 12 годин з моменту ув'язнення осіб, які не вчинили важких злочинів. При цьому особи, заарештовані за дрібні правопорушення, наприклад, за домашнє насильство або ті, хто не з'явився до суду більше двох разів протягом останнього року, під автоматичне звільнення не підпадають.

Особи, арештовані за тяжкі злочини та кримінальні проступки, які не підлягають автоматичному звільненню, мають оцінюватися на предмет ризику вчинення нового злочину або неявки до суду у разі звільнення. Автори ініціативи пропонували, що співробітники, які проводять оцінку, збиратимуть певну інформацію, включаючи рівень ризику кожної особи, визначений за допомогою інструменту досудової оцінки ризиків. Як правило, співробітники повинні звільняти осіб, які мають низький рівень ризику. Залежно від правил, встановлених кожним судом першої інстанції, певні особи із середнім ступенем ризику також можуть бути звільнені персоналом з оцінки або суддею. Звільнені особи можуть бути зобов'язані дотримуватися певних умов, наприклад, регулярно реєструватися в окружному центрі пробації або у системі електронного моніторингу. Однак умови для осіб з низьким рівнем ризику можуть не включати нагляд. Суд міг змінити умови з поважних причин. На відміну від чинного законодавства, жодні платежі не могли стягуватися як умова звільнення. Особи з високим та середнім ступенем

ризик, які не звільняються, та деякі інші (наприклад, обвинувачені у певних тяжких злочинах, включаючи вбивство або підпал будинку) залишатимуться в окружній в'язниці до пред'явлення обвинувачення. Оцінка і будь-яке звільнення повинні бути завершені не пізніше, ніж через 36 годин з моменту поміщення особи до в'язниці.

Ініціатива покладала на суди штату першої інстанції відповідальність за досудове оцінювання, зокрема: 1) визначення рівня ризику за допомогою інструментів досудової оцінки ризиків; 2) збір додаткової інформації, пов'язаної з ризиком особи; 3) звільнення певних осіб на основі рівня їхнього ризику та 4) пропонування суду умов досудового звільнення. Суд першої інстанції може залучати працівників суду для проведення оцінки або укладати контракти з певними місцевими установами штату (наприклад, з окружним відділом пробачії) для виконання цих завдань. Якщо ні суд, ні наявна місцева установа штату не бажають або не можуть це робити, то суд може укласти договір з новою місцевою установою, створеною спеціально для виконання цих функцій.

Ініціатори поправки переконували, що вона зменшить кількість ув'язнених в окружних в'язницях, оскільки більше осіб буде звільнено з-під варті на досудовому слідстві замість того, щоб залишатися у в'язниці. Наприклад, деякі люди, які не змогли б сплатити заставу, будуть звільнені за новим досудовим процесом.

Аргументи на підтримку ініціативи зводилися до таких наступних: нівелювання впливу майнового чинника, оскільки за відсутності фінансових можливостей сплатити заставу, особа залишається у тюрмі; посилення справедливості: запровадження комп'ютеризованої оцінки сприятиме більшій справедливості, оскільки рішення буде прийматися виходячи зі ступеня небезпечності особи для суспільства, а не її майнового статусу; економія коштів; захист невинних: чинна система грошової застави може змусити невинних людей визнати себе винними у злочинах, яких вони не скоювали.

Противники прийняття поправки наполягали на тому, що вона: небезпечна, оскільки скасовує можливість внесення застави для кожного каліфорнійця і замінює це право на систему комп'ютерного профілювання, що адмініструється округом; дискримінує меншини та мешканців районів з високою концентрацією іммігрантів

та мешканців з низьким рівнем доходу; збільшить навантаження на суди першої інстанції штатів, а також на окружних прокурорів і державних захисників; затратна, оскільки вимагатиме проведення додаткових судових слухань для скасування рішення комп'ютера, що призведе до додаткових затримок в і без того перевантажених судах; позбавляє права на заставу.

На кампанію проти цієї ініціативи було витрачено понад \$10 млн і зрештою вона не була підтримана: проти проголосувало 56,41% виборців, а «за» – лише 43,59% [6]. При цьому на референдумі були підтримані такі прогресивні ініціативи, як фінансування досліджень стовбурових клітин, надання права голосу достроково звільненим злочинцям тощо. Це був несподіваний результат для більшої частини політичної еліти штату, враховуючи статус Каліфорнії як прогресивного штату, де грошова застава вже давно є непопулярною практикою.

Важливість позитивного ставлення громадськості до використання інструментів ШІ у прийнятті політичних рішень зумовлена трьома чинниками. По-перше, з нормативного погляду, громадськість є основною зацікавленою стороною у формуванні майбутнього ШІ, а отже, її слід залучати до дискусій щодо управління ШІ. По-друге, як і у випадку з іншими технологіями (наприклад, ядерною енергією та генетично модифікованими продуктами харчування, вакцинами тощо) громадськість може впливати на розробку та впровадження ШІ. Розуміння того, що вони думають про ШІ, допоможе нам передбачити майбутні політичні суперечки та поведінку споживачів. По-третє, це питання легітимності: якщо громадяни не довірятимуть рішенням, прийнятим з використанням ШІ, то вони й не будуть їх виконувати.

Навіть попри те, що системи ШІ дедалі ширше впроваджуються в суспільство, більшість наявних праць з етики ШІ зосереджуються на етичних принципах, розроблених технологічними компаніями, урядами, аналітичними центрами або академічними установами [7]. Значна частина робіт має рекомендаційний характер: вони описують, як мають виглядати ідеальні «етичні» системи ШІ. Ці принципи, такі як повага до прав людини або збереження приватності, спрямовані на захист суспільного добробуту, але участь громадськості у формуванні цих етичних принципів була обмеженою або надто затратною.

Міжнародний проєкт Moral Machine виявив, що респонденти в різних регіонах і культурах мають різні уподобання щодо поведінки автономних автомобілів у моральних дилемах [8]. Такі дослідження неминуче ускладнюють етику ШІ, оскільки виявляють, що споживачі та виборці не згодні з тим, як слід розробляти та впроваджувати системи ШІ.

Розуміння суспільних настроїв може допомогти політикам та чиновникам передбачити масову мобілізацію навколо питань, пов'язаних зі ШІ, особливо навколо закликів заборонити конкретні програми з його використанням. По-перше, вони виступають як безпосередні споживачі і бойкотують продукти або послуги, вироблені за допомогою цієї технології. Наприклад, європейські споживачі сприймають ГМ-продукти як ризиковані і неохоче їх споживають. По-друге, громадськість може вимагати змін у регулюванні шляхом масової мобілізації, активізму та голосування. Виборці, які виступають проти атомної енергетики, голосували на референдумах про її заборону (Австрія в 1978 році та Італія у 2011 році) або поступову відмову від використання атомних станцій (Швейцарія у 2017 році).

Термін «надійний ШІ» є досить популярним у заявах з етики ШІ, опублікованих урядами, технологічними компаніями та групами громадянського суспільства. У найбільш узагальненому вигляді серед принципів надійного ШІ називають доброчинність, незлочинність, автономію, справедливість і розумілість [9]. Ажіотаж навколо надійного ШІ зумовив появу низки праць, присвячених розробці алгоритмів та інституцій, яким довірятиме громадськість. Головний недолік цих робіт полягає в тому, що вони не приділяють належної уваги суб'єктивному сприйняттю та досвіду користувачів, на яких впливають системи ШІ. Водночас дослідження громадської думки не встигають за цими теоретичними напрацюваннями; натомість вони зосереджуються на загальному ставленні громадськості до ШІ, відірваному від будь-якого технічного чи інституційного контексту.

Наукові розвідки про надійний ШІ часто пропонують соціотехнічні рамки для побудови систем ШІ, яким довірятиме громадськість. Ці концепції зазвичай передбачають створення технічних або інституційних особливостей, що сприяють безпеці, справедливості та прозорості. Деякі з запропонованих рішень включають

пояснення того, як працює система ШІ, створення документації про процес розробки, вимогу проведення аудиту третьою стороною, винагороду для тих, хто виявляє недоліки в системах ШІ, створення бази даних про інциденти, пов'язані зі штучним інтелектом, і прийняття нормативно-правових актів тощо [10]. У подальшому ці пропозиції знайшли своє відображення у низці нормативно-правових актів, що регулюють використання ШІ у різних країнах світу.

Існує досить широкий масив даних щодо ставлення громадян до таких основних сфер застосування ШІ та алгоритмів, як розпізнавання облич, алгоритми персоналізації та автоматизація робочих місць.

Технологія розпізнавання облич та алгоритми персоналізації стали звичними, але останнім часом викликають значну увагу громадськості. Алгоритми розпізнавання облич, що використовуються для ідентифікації, верифікації та класифікації осіб на основі їхніх рис обличчя, станом на 2020 рік були розгорнуті або схвалені до розгортання у 109 країнах [11]. Ця технологія є стандартною у різноманітних додатках, таких як розблокування смартфонів або позначення людей на фотографіях; зараз правоохоронні органи, роботодавці та бізнес все частіше вдаються до цієї технології. Однак така технологія становить ризик для приватності та громадянських свобод. Навіть якщо алгоритми розпізнавання облич стануть більш точними, критики стверджують, що ця технологія збільшить можливості правоохоронних органів, урядів і навіть приватних компаній контролювати суспільство, завдаючи непропорційно великої шкоди маргіналізованим групам [12].

Ставлення громадськості до технології розпізнавання облич має свої особливості, хоча деякі ключові тенденції повторюються в різних опитуваннях. По-перше, в деяких країнах громадськість більш прихильно ставиться до технології розпізнавання облич, що використовується правоохоронними органами, ніж до приватних суб'єктів. По-друге, підтримка корелює з демографічними змінними, такими як країна проживання, раса та політичні уподобання. Демографічні змінні, включаючи країну проживання, расу та політичні погляди, корелюють з підтримкою технології розпізнавання облич. Так респонденти в Китаї висловили найвищу підтримку використання розпізнавання облич (67%), а 38% дорослих у Німеччині, 50% у Великій Британії та 48% у США підтримують

використання цієї технології [13]. При цьому респонденти в Китаї, порівняно з іншими країнами, вважають технологію розпізнавання облич більш зручною та ефективною, а також менш ризикованою з точки зору конфіденційності, дискримінації та стеження. Цей висновок, мабуть, не є дивним з огляду на поширеність технології, що використовується правоохоронними органами та приватними особами в Китаї.

Алгоритми персоналізації використовують дані користувача в Інтернеті та поза ним для створення онлайн-контенту, відповідно до його уподобань. Алгоритми персоналізації застосовуються, зокрема, в рекомендаційних системах для новинних статей у соціальних мережах, таргетованій рекламі в Інтернеті або індивідуальному ціноутворенні. Хоча алгоритми персоналізації стали повсюдними, однак критикуються низкою дослідників. З погляду приватності, персональні та поведінкові дані користувачів збираються без їхньої інформованої згоди технологічними компаніями для продажу рекламодавцям [14]. До того ж персоналізація може посилювати погляди людей з екстремістськими політичними переконаннями [15].

При цьому, хоча громадськість може не до кінця розуміти, як працюють алгоритми персоналізації, вона виступає проти збору та використання певних типів даних або проти певних видів персоналізації. Дані національних репрезентативних опитувань у Німеччині, Великій Британії та США показали, що населення цих країн виступає проти того, щоб технологічні компанії збирали чутливу інформацію, як-от особисті трагедії чи доходи домогосподарств; крім того, респонденти висловили консенсус проти використання персоналізації в політичній агітації [16]. Зрештою користувачі змирилися з тим, що алгоритми персоналізації є дуже непрозорими, а збір даних неодмінною умовою користування сучасними комунікаційними технологіями.

Що стосується автоматизації робочих місць, то у цьому сенсі показовими є результати опитувань громадян Євросоюзу, проведене Євробарометром у кінці 2024 року. Воно з'ясувало, що половина європейців вважають, що новітні цифрові технології, включаючи штучний інтелект, позитивно впливають на: роботу (66%); на економіку, якість їхнього життя та соціальне забезпечення (62%); на суспільство (56%) [17]. Проте частка тих, хто вважає, що новітні

технології позитивно впливають на економіку (-13%), суспільство (-8%.) та якість життя (-5%), зменшилася порівняно з березнем 2017 року.

Має місце залежність від соціально-демографічних характеристик опитаних. Так молодь, респонденти з вищою освітою, ті, у кого немає фінансових труднощів, та працюючі частіше вважають, що ці технології мають позитивний вплив.

При цьому переважна більшість опитаних вважає, що правила, які регулюють ризики та максимізації переваг цифрових технологій на робочому місці є досить важливими. Зокрема 82% респондентів заявили, що захист приватності працівників важливий для зменшення ризиків і максимізації переваг цифрових технологій, зокрема штучного інтелекту, на робочому місці. Також особливе значення має залучення працівників та їхніх представників до розробки та впровадження нових технологій (77%), забезпечення більшої прозорості у використанні цифрових технологій для прийняття кадрових рішень (75%), заборона повністю автоматизованих процесів прийняття рішень (74%) та обмеження автоматизованого моніторингу працівників (72%).

Загалом більшість європейців позитивно ставляться до робіт та ІІІ на робочому місці, але висловлюють занепокоєння щодо їхнього негативного впливу на працевлаштування. Хоча більшість (62%) позитивно сприймають використання робіт і ІІІ на робочому місці, проте майже третина (32%), ставляться до цього негативно. Однак це сприйняття суттєво відрізняється залежно від країни: 86% респондентів у Данії заявили, що ставляться до цього позитивно, порівняно з менш ніж половиною в Португалії та Греції (по 48%). Більшість опитаних погоджується з тим, що роботи та ІІІ підвищують швидкість виконання завдань (73%), що ІІІ необхідний, оскільки він може виконувати роботу, яка вважається нудною або повторюваною (66%), і що роботи та ІІІ можуть використовуватися для прийняття точних рішень на робочому місці (53%). Водночас 66% вважають, що у результаті широкомасштабного їх застосування зникне більше робочих місць, ніж буде створено нових, іншими словами вони забирають роботу у людей. До того ж більше як половина (61%) вважають, що ці технологічні удосконалення негативно впливають на комунікацію між колегами.

Переважна більшість вважає, що роботи та ШІ мають позитивний вплив на суспільство, але слід ретельно контролювати їх використання. Так 63% опитаних вважають, що вони інтелект допомагають людям виконувати свою роботу або щоденні завдання вдома, хоча 84% вказують на необхідність прискіпливого нагляду за ними.

Використання цифрових технологій, в тому числі ШІ, для багатьох видів діяльності на робочому місці розглядається в негативному світлі. Загалом респонденти занепокоєні використанням цифрових технологій, зокрема ШІ, для автоматичного звільнення працівників (78%), моніторингу працівників (63%), оцінки ефективності роботи працівників (57%), відбору кандидатів на роботу (57%), збору додаткової інформації про кандидатів на роботу (50%), а також для збору, обробки та зберігання персональних даних працівників (49%). Однак близько двох третин (67%) позитивно сприймають використання таких технологій для підвищення безпеки працівників, а відносна більшість позитивно сприймає їх використання для розподілу завдань між працівниками або управління їхніми робочими графіками та змінами (49% «позитивно» проти 44% «негативно»). Позитивне сприйняття використання цих технологій на робочому місці більш поширене серед молодших респондентів, респондентів з вищою освітою та тих, хто зараз працює.

Отже, важливість позитивного ставлення громадськості до використання інструментів ШІ у прийнятті політичних рішень зумовлена трьома чинниками. По-перше, громадськість є основною зацікавленою стороною у формуванні майбутнього ШІ, а отже, її слід залучати до дискусій щодо управління ШІ. По-друге, як і у випадку з іншими технологічними нововведеннями громадськість може впливати на розробку та впровадження ШІ. Розуміння того, що вони думають про ШІ, допоможе нам передбачити майбутні політичні суперечки та поведінку споживачів. По-третє, це питання легітимності: якщо громадяни не довірятимуть рішенням, прийнятим з використанням ШІ, то вони й не будуть їх виконувати.

Розуміння ставлення громадян до нововведень може допомогти політикам та чиновникам передбачити масову мобілізацію навколо питань, пов'язаних із використанням ШІ у прийнятті політичних або адміністративних рішень. Адже громадяни є безпосередніми споживачами, які можуть вдатися до бойкоту продуктів або послуг, що надаються за допомогою цієї технології. Крім того, громад-

ськість може вимагати змін у нормативно-правовому регулюванні розробки та застосування ІІІ шляхом масової мобілізації, активізму та голосування, що може створити перешкоди для розвитку технологічного сектора в країні.

1. Stevenson M. Assessing risk assessment in action. *Minnesota Law Review*. 2018. Vol. 103:303.
2. Bates D.W., Saria S., Ohno-Machado L., Shah A., Escobar G. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Affairs*. 2014. Vol. 33. № 7. P. 1123–1131.
3. Angwin J., Larson J., Mattu S., Kirchner L. Machine bias. ProPublica. 2016 URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
4. Obermeyer Z., Powers B., Vogeli C., Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*. 2019. Vol. 366(6464). P. 447–453.
5. Official voter information guide, California general election 2020. URL: <https://vig.cdn.sos.ca.gov/2020/general/pdf/complete-vig.pdf>
6. General Election – Statement of Vote, November 3, 2020. URL: <https://www.sos.ca.gov/elections/prior-elections/statewide-election-results/general-election-november-3-2020/statement-vote>
7. Fjeld J., Achten N., Hilligoss H., Nagy A., Srikumar, M. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for AI. Berkman Klein Center for Internet & Society, 2020. URL: <https://ssrn.com/abstract=3518482>
8. Awad E., Dsouza S., Kim R. The moral machine experiment. *Nature*. 2018. Vol. 563(7729). P. 59–64.
9. Thiebes S., Lins S., Sunyaev A. Trustworthy artificial intelligence. *Electronic Markets*. 2021. Vol. 31. P. 447–464. URL: <https://doi.org/10.1007/s12525-020-00441-4>.
10. Toward trustworthy AI development: Mechanisms for supporting verifiable claims (2020). URL: <https://cdn.governance.ai/Toward-Trustworthy-AI-Development.pdf>
11. The facial recognition world map. Surfshark, 2020. URL: <https://perma.cc/6HUK-9Q4D>
12. West S.M., Whittaker M., Crawford K. Discriminating systems: Gender, race, and power in AI. AI New Institute, 2018. URL: <https://ainowinstitute.org/publications/discriminating-systems-gender-race-and-power-in-ai-2>
13. Kostka G., Steinacker L., Meckel M. Between security and convenience: Facial recognition technology in the eyes of citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding of Science*. 2021. Vol. 30. № 6. P. 671–690.
14. Zuboff S. The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs, 2019.
15. Stark B., Stegmann D., Magin, M., Jurgens P. Are algorithms a threat to democracy? the rise of intermediaries: A challenge for public discourse. AlgorithmWatch, 2020. URL: <https://perma.cc/WS6G-NQ5U>
16. Kozyreva A., Lorenz-Spreen P., Hertwig R., Lewandowsky S., Herzog S. M. Public attitudes towards algorithmic personalization and use of personal data online: Evidence from Germany, Great Britain, and the United States. *Humanities and Social Sciences Communications*. 2021. Vol. 8. № 1.

P. 1–11. 17. Artificial Intelligence and the future of work. Eurobarometer, 2025. URL: <https://europa.eu/eurobarometer/surveys/detail/3222>

Vitushko Maksym. The influence of public opinion on the use of artificial intelligence in political decision-making

The increasing use of artificial intelligence and algorithms in decision-making raises a number of issues. First and foremost, we are talking about algorithmic bias and lack of transparency in the process. The author uses the example of the 2020 referendum in California on the introduction of algorithmic assessment when deciding whether to release a person from bail.

The importance of a positive public attitude towards the use of AI tools in policy decision-making is driven by three factors. First, from a regulatory perspective, the public is a key stakeholder in shaping the future of AI and should therefore be involved in discussions about AI governance. Second, as with other technologies (e.g. nuclear energy and genetically modified foods, vaccines, etc.), the public can influence the development and implementation of AI. Understanding what they think about AI will help us predict future political controversies and consumer behaviour. Thirdly, there is the issue of legitimacy: if citizens do not trust decisions made using AI, they will not follow them.

Understanding public sentiment can help politicians and officials anticipate mass mobilisation around AI-related issues, especially around calls to ban specific AI applications. Firstly, they act as direct consumers and boycott products or services produced by this technology. For example, European consumers perceive GM foods as risky and are reluctant to consume them. Second, the public can demand regulatory changes through mass mobilisation, activism and voting. Voters opposed to nuclear power have voted in referendums to ban it (Austria in 1978 and Italy in 2011) or phase out nuclear power plants (Switzerland in 2017).

Three key areas where the use of AI will have the greatest impact on public attitudes are considered, including face recognition, personalisation algorithms, and workplace automation. Based on the analysis of the results of public opinion polls, the factors influencing the attitude of citizens towards AI are identified.

Understanding citizens' attitudes to innovations can help politicians and officials anticipate mass mobilisation around issues related to the use of AI in political or administrative decision-making. After all, citizens are direct consumers who may resort to boycotting products or services provided by this technology. In addition, the public can demand changes in the legal and regulatory framework for the development and use of AI through mass mobilisation, activism and voting, which can create obstacles to the development of the technology sector in the country.

Key words: artificial intelligence, legitimacy, democracy, face recognition, algorithmic bias, automation.