

ПРАВИЛНО ИЗГРАЖДАНЕ НА БЪДЕЩЕТО

ИЗКУСТВЕН ИНТЕЛЕКТ И ОСНОВНИ ПРАВА

РЕЗЮМЕ



3

Основни констатации и становища на FRA

5

Защита на основните права — обхват,
оценки на въздействието и отчетност

10

Недопускане на дискриминация, защита
на личните данни и достъп до правосъдие:
три хоризонтални теми

© Агенция на Европейския съюз за основните права, 2021 г.

Възпроизвеждането е разрешено, при условие че се посочи източникът.

За използването или възпроизвеждането на снимки или други материали, за които Агенцията на Европейския съюз за основните права, не е носител на авторското право, трябва да се иска разрешение направо от носителите на това право.

Нито Агенцията на Европейския съюз за основните права, нито което и да е лице, действащо от името на Агенцията, носят отговорност за начина, по който би могла да бъде използвана съдържащата се в настоящата публикация информация.

Люксембург: Служба за публикации на Европейския съюз, 2021 г.

Print	ISBN 978-92-9461-220-5	doi:10.2811/975264	TK-04-20-657-BG-C
PDF	ISBN 978-92-9461-247-2	doi:10.2811/399522	TK-04-20-657-BG-N

Снимки:

Корица: © HQUALITY/Adobe Stock

Страница 1: © Mykola Mazuryk/AdobeStock

Страница 3: © Mimi Potter/AdobeStock

Страница 4: © Sikov/Adobe Stock

Страница 7: © Gorodenkoff/Adobe Stock

Страница 9: © VideoFlow/Adobe Stock

Страница 11: © Monsitj/Adobe Stock

„ИИ е чудесно нещо, но трябва да се научим как да го използваме.“
(Частно дружество, Испания)

„Съществува риск да започнем да се доверяваме прекалено много на машината.“
(Държавна администрация, Франция)

Изкуственият интелект (ИИ) е все по-често използван в частния и публичния сектор, което влияе върху ежедневието. Някои хора считат ИИ за края на човешкия контрол върху машините. Други го приемат като технология, която ще помогне на човечеството да се справи с някои от най-неотложните предизвикателства. Въпреки че нито една от двете представи не е напълно точна, опасенията относно въздействието на ИИ върху основните права очевидно се засилват, което оправдава усилието на правозащитните организации за контрола върху използването му.

В доклада на FRA „Правилно изграждане на бъдещето – изкуствен интелект и основни права“ е представен обзор на текущото използване в ЕС на технологии, свързани с ИИ, и се анализира отражението им върху основните права. Разгледани са случаи на употреба в четири основни области: социални помощи, прогнозиране в полицейската област, здравни услуги и целево рекламиране.

В настоящото резюме се представят основните гледни точки, описани в доклада.

Какво представлява ИИ

Няма общоприето определение за ИИ. Понятието не се отнася за конкретни приложения, а отразява последните технологични разработки, които обхващат различни технологии.

В проучването на FRA не се използва строго определение за ИИ в случаите на употреба, представени в основния доклад. При интервютата е дадено определение за ИИ в широк смисъл, като е направена препратка към определението на Експертната група на високо равнище на Европейската комисия по въпросите на изкуствения интелект (AI HLEG):

„Понятието изкуствен интелект (ИИ) се отнася за системи, които показват интелигентно поведение, като анализират своята среда и предприемат действия – с известна степен на самостоятелност – за да постигат конкретни цели. Системите на основата на ИИ може да са изцяло софтуерно базирани и да действат във виртуалния свят (напр. гласови асистенти, софтуер за анализ на изображения, търсачки, системи за разпознаване на речта и лицето) или ИИ да бъде вграден в хардуерни устройства (напр. усъвършенствани роботи, автономни автомобили, дронове или приложения за „интернет на предметите“).“

*Това първоначално определение на AI HLEG беше обект на допълнително обсъждане в групата. Вж. AI HLEG (2019), **A definition of AI: Main capabilities and disciplines (Определение за ИИ: основни възможности и дисциплини)**.*

Какъв е обхватът на проучването?

FRA проведе проучване на място в пет държави — членки на ЕС: Естония, Финландия, Франция, Нидерландия и Испания. Агенцията събра информация от предприятия и институции, занимаващи се с проектиране и използване на системи с ИИ в ключови частни и държавни сектори, относно начина, по който разглеждат въпроси, свързани с основните права.

Проучването, което се основава на 91 индивидуални интервюта, събра информация относно:

- целта и практическото приложение на технологиите с ИИ;
- извършените оценки при използване на ИИ и приложимата нормативна уредба и механизми за надзор;
- осведомеността за проблемите с основните права и евентуални предпазни мерки; и
- плановете за бъдещето.

В допълнение са интервюирани 10 експерти, които се занимават с контрола или наблюдението на потенциални нарушения на основните права, свързани с използването на ИИ, включително представители на гражданското общество, юристи и надзорни органи.

За по-подробно описание на методологията на проучване и задаваните въпроси по време на интервютата вижте Приложение 1 към основния доклад, достъпно на [уебсайта на FRA](#).

„Най-важно е да се справяме по-ефективно със случаите. За да се случи това, работната сила, или хората, обработващи случаите, трябва да се използва по възможно най-ефективен начин.“

(Държавна администрация, Нидерландия)

„Когато изпитвахме системата, не обърнахме голямо внимание на правните аспекти, а дали системата е рентабилна.“

(Частно дружество, Естония)

Работата на FRA в областта на ИИ, големите информационни масиви и основните права

Докладът на FRA относно ИИ и основните права е основната публикация, съставена в резултат от **проекта на FRA в областта на изкуствения интелект, големите информационни масиви и основните права**. Целта на проекта е да оцени положителните и отрицателните последици от новите технологии, включително ИИ и големите информационни масиви, върху основните права.

Докладът е изготвен въз основа на констатациите от няколко предишни доклада:

- **Facial recognition technology: fundamental rights considerations in the context of law enforcement** („Технология за лицево разпознаване: съображения, свързани с основните права, в контекста на правоприлагането“) (2019 г.): този доклад очертава и анализира възникналите предизвикателства за основните права, при използване на технология за лицево разпознаване в реално време от държавните органи за целите на правоприлагането. Освен това той представя накратко стъпките, които да се предприемат, за да се избегне нарушаването на права.
- **Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights** („Качество на данните и изкуствен интелект — намаляване на грешките и отклоненията с цел защита на основните права“) (2019 г.): този доклад маркира значението на осведомеността и избягването на лошо качество на данните.
- **#BigData: Discrimination in data-supported decision making** („#BigData: дискриминация при вземане на решения, подкрепени от данни“) (2018 г.): този специален документ разглежда случаите, където може да възникне такава дискриминация, и предлага възможни решения.

Като част от проекта, FRA разглежда също възможността за проучване на конкретни примери за предизвикателства, свързани с основните права, при използване на алгоритми за вземане на решения, чрез онлайн експерименти или симулационни изследвания.

Основни констатации и становища на FRA

Новите технологии промениха съществено начина, по който организираме и живеем живота си. По-специално, новите технологии, основани на данни, насърчиха развитието на изкуствения интелект (ИИ), включително увеличената автоматизация на задачи, които обикновено се изпълняват от хора. Здравната криза, предизвикана от COVID-19, стимулира използването на ИИ и обмяна на данни, като така се създадоха нови възможности, но възникнаха също предизвикателства и заплахи за правата на човека и основните права.



Разработките в областта на ИИ получиха широк отзвук сред медиите, гражданското общество, академичните среди, органите за защита на правата на човека и създателите на политики. Голяма част от този отзвук акцентира върху потенциала на изкуствения интелект да подкрепи икономическия растеж. По-малко внимание се обръща на начина, по който различните технологии биха могли да засегнат основните права. Към момента все още не разполагаме с голям обем емпирични доказателства за широкия набор от права, върху които влияе ИИ, или за необходимите предпазни мерки, за да се гарантира, че практическото използване на ИИ е в съответствие с основните права.

На 19 февруари 2020 г. Европейската комисия публикува Бяла книга за изкуствения интелект – *Европа в търсене на високи постижения и атмосфера на доверие*. Тя описва основните принципи на бъдеща регулаторна рамка за ИИ в Европа. Бялата книга отбелязва, че е важно тази рамка да бъде обвързана с основните ценности на ЕС, включително зачитане на правата на човека – член 2 от Договора за Европейския съюз (ДЕС).

Докладът на FRA за ИИ и основните права подкрепя тази цел, като анализира отражението върху основните права при използването на ИИ. Въз основа на конкретни „случаи на употреба“ на ИИ в избрани области, в доклада се обръща внимание на реалната ситуация с предизвикателствата и възможностите при използването на ИИ, свързани с основните права.

Докладът е съставен въз основа на 91 интервюта с длъжностни лица от държавната администрация и персонал от частни дружества в избрани държави – членки на ЕС. На участниците са зададени въпроси относно използването на ИИ, осведомеността им за свързаните с това проблеми с основните права и практиките при оценката и намаляването на рисковете, свързани с използването на ИИ.

В допълнение са проведени 10 интервюта с експерти, чиято работа по някакъв начин е свързана с потенциалните предизвикателства на ИИ по отношение на правата на човека. Тази група включва държавни органи (напр. контролни и надзорни органи), неправителствени организации и юристи.



Нормативна уредба

Всеобщата рамка за основните права (*), която се прилага за използването на ИИ в ЕС, включва Хартата на основните права на ЕС (Хартата) и Европейската конвенция за правата на човека.

Съществуват и много други относими инструменти на Съвета на Европа и международни актове в областта на правата на човека. Те включват Всеобщата декларация за правата на човека от 1948 г. и значимите конвенции на ООН за правата на човека (**).

В допълнение защитата на основните права в контекста на ИИ се подкрепя от специфичното за секторите вторично право на ЕС, по-конкретно *достиженията на правото* на ЕС в областта на защитата на личните данни и законодателството на ЕС за недопускане на дискриминация. На последно място, в тази област се прилага също националното законодателство на държавите — членки на ЕС.

(*) За повече информация вж. FRA (2012), *Bringing rights to life: The fundamental rights landscape of the European Union („Прилагане на правата: ситуацията в Европейския съюз с основните права“)*, Люксембург, Служба за публикации на Европейския съюз.

(**) Тези значими конвенции включват: Международния пакт за граждански и политически права от 1966 г.; Международния пакт за икономически, социални и културни права от 1966 г.; Международната конвенция за премахване на всички форми на расова дискриминация от 1965 г.; Конвенция за премахване на всички форми на дискриминация по отношение на жените от 1979 г.; Конвенцията против изтезанията от 1984 г.; Конвенцията за правата на детето от 1989 г.; Конвенцията за правата на хората с увреждания от 2006 г. и Международната конвенция за защита на всички лица от насилствено изчезване от 2006 г.

За повече информация относно всеобщата международна правна рамка в областта на правата на човека, включително механизмите за прилагане, вж. напр. De Schutter, O. (2015), *International Human Rights Law: Cases, Materials, Commentary*, Cambridge, Cambridge University Press, 2nd edition.

ЗАЩИТА НА ОСНОВНИТЕ ПРАВА — ОБХВАТ, ОЦЕНКИ НА ВЪЗДЕЙСТВИЕТО И ОТЧЕТНОСТ

Относно пълния обхват на основните права във връзка с ИИ

Използването на системи с ИИ засяга широк набор от основни права, независимо от областта на приложение. Те включват следните, но не само: неприкосновеност на личния живот, защита на личните данни, недопускане на дискриминация и достъп до правосъдие.

Хартата на основните права на ЕС (Хартата) придобива правнообвързващ характер през декември 2009 г. и има същата правна сила като договорите за ЕС. Тя обединява гражданските, политическите, икономическите и социалните права в един текст. Съгласно член 51, параграф 1 от Хартата, институциите, органите, службите и агенциите на Съюза трябва да зачитат всички права, описани в Хартата. Държавите — членки на ЕС, следва да изпълняват това изискване, когато прилагат правото на Съюза. Това се отнася в еднаква степен за ИИ, както за всяка друга област.

Работата на място по това проучване показва, че с наименованието „ИИ“ се обозначава голямо разнообразие от системи. Анализираните технологии включват различни нива на автоматизация и сложност. Те се различават също по отношение на своя мащаб и потенциално въздействие върху хората.

Констатациите на FRA показват, че използването на системи с ИИ дава отражение върху широк спектър от основни права, независимо от областта на приложение. Те включват следните, но не само: неприкосновеност на личния живот, защита на личните данни, недопускане на дискриминация и достъп до правосъдие. Въпреки това при разглеждане на въздействието на ИИ върху основните права интервютата показват, че обхватът често е ограничен до конкретни права.

Необходимо е да се разгледа по-широк спектър от права при използване на ИИ в зависимост от технологията и областта на използване. Освен правата, свързани с неприкосновеността на личния живот и защитата на личните данни, равнопоставеността и недопускането на дискриминация и достъпа до правосъдие, може да се обмисли разглеждането и на други права. Например човешкото достойнство, правото на обществено осигуряване и социално подпомагане, правото на добра администрация (приложимо най-вече за държавния сектор) и защитата на потребителите (от особено значение за предприятията). В зависимост от контекста на използване на ИИ е необходимо разглеждане и на всяко друго право, защитено от Хартата.

СТАНОВИЩЕ 1 НА FRA

При въвеждане на нови политики и приемане на ново законодателство в областта на ИИ законодателят на Съюза и държавите членки, действащи в рамките на обхвата на правото на Съюза, трябва да гарантират, че е взето под внимание зачитането на пълния спектър от основни права, заложен в Хартата и Договорите за ЕС. Съответните политики и закони е необходимо да бъдат придружени от мерки за защита на конкретни основни права.

При това ЕС и държавите членки следва да се позовават на надеждни доказателства за въздействието на ИИ върху основните права, за да гарантират, че при всяко ограничение на определени основни права се зачитат принципите на необходимост и пропорционалност.

Необходимо е законът да предвижда съответните мерки за ефективна защита срещу произволна намеса в основните права и да предоставя правна сигурност както за работодателите, така и за потребителите на ИИ. Доброволните схеми за спазване и защита на основните права при разработването и използването на ИИ могат допълнително да помогнат за намаляване на нарушенията на правата. В съответствие с минималните изисквания за юридическа яснота, като основен принцип на върховенството на закона и предпоставка за гарантиране на основните права, законодателят трябва да отдаде дължимото внимание при определяне на обхвата на всяко такова законодателство в областта на ИИ.

Предвид разнообразните технологии, включени в понятието ИИ, и липсата на знания за пълния обхват на потенциалното въздействие върху основните права, е необходима редовна оценка на правните определения на понятията, свързани с ИИ.

СТАНОВИЩЕ 2 НА FRA

Законодателят на Съюза трябва да разгледа възможността за извършване на задължителни оценки на въздействието, които обхващат пълния спектър от основните права. Те трябва да включват частния и държавния сектор и да се прилагат преди използването на системи с ИИ. Оценките на въздействието трябва да отчетат различното естество и обхват на технологиите с ИИ, включително нивото на автоматизация и сложност, както и потенциалните вреди. Те трябва да включват основни изисквания за преценка, които могат да послужат и за повишаване на осведомеността за потенциалните отражения върху основните права.

Оценките на въздействието трябва да почиват на установената добра практика в други области и да се повтарят редовно по време на внедряването, ако е целесъобразно. Те трябва да се извършват по прозрачен начин. Резултатите и препоръките от тях трябва, доколкото е възможно, да бъдат направени обществено достояние. За да подпомогнат процеса на оценка на въздействието, от предприятията и държавните администрации трябва да се изисква да събират необходимата информация за задълбочена оценка на потенциалното въздействие върху основните права.

ЕС и държавите членки трябва да обмислят прилагането на целенасочени действия в подкрепа на предприятията и институциите, които разработват, използват или планират да използват системи с ИИ, за да се гарантира ефективно спазване на задълженията им за оценка на въздействието върху основните права. Тези действия може да включват финансиране, насоки, обучения или повишаване на осведомеността. Те трябва да са насочени най-вече, но не само, към частния сектор.

ЕС и държавите членки трябва да разгледат възможността за използване на съществуващите инструменти, например контролни списъци или инструменти за самооценка, разработени на европейско и международно равнище. Това включва разработените от групата на високо равнище на ЕС инструменти, свързани с изкуствения интелект.

Използване на ефективна оценка на въздействието за предотвратяване на неблагоприятни ефекти

Предварителните оценки на въздействието се концентрират главно върху техническите въпроси. Те рядко разглеждат потенциалните ефекти върху основните права. Това е така, защото липсват знания за начина, по който ИИ засяга тези права.

Внедряването на системи с ИИ засяга широк спектър от основни права, независимо от областта на приложение. Съгласно член 51, параграф 1 от Хартата държавите — членки на ЕС, трябва да зачитат всички права, описани в Хартата, при прилагане на правото на Съюза. Съгласно съществуващите международни стандарти, по-специално Ръководните принципи на ООН за бизнеса и правата на човека (UNGP), бизнесът трябва да има въведен „процес на надлежна проверка на спазването на правата на човека, за да може да идентифицира, предотвратява, намалява и отчита начина, по който преодолява въздействието си върху правата на човека“ (принципи 15 и 17). Това не трябва да зависи от размера на предприятието и сектора, и включва дейностите, работещи с ИИ.

Докато следва ангажиментите си към UNGP, ЕС прие няколко законодателни акта, насочени към специфични за сектора инструменти, по-специално в контекста на задълженията, свързани с надлежна проверка за спазването на правата на човека. В момента текат обсъждания за предлагане на ново вторично законодателство на ЕС. Това законодателство ще изисква от предприятията да извършва надлежна проверка на потенциалното въздействие на своите операции и вериги за доставки върху правата на човека и околната среда. Вероятно това законодателство ще бъде междусекторно и ще предвижда санкции за несъответствие, които би трябвало да обхващат и използването на ИИ. Вижте един от последните доклади на FRA със заглавие *Business and Human rights – access to remedy („Бизнесът и правата на човека — достъп до средства за правна защита“)*, който призовава за по-добри хоризонтални правила за надлежна проверка за спазването на правата на човека за дружества, установени в ЕС.

Оценките на въздействието са важен инструмент както за предприятията, така и за държавната администрация, за да се намали потенциалното неблагоприятно въздействие от дейностите им върху основните права. Правото на Съюза в определени сектори изисква някои форми на оценка на въздействието, например Оценка на въздействието върху защитата на данните съгласно Общия регламент относно защитата на данните (ОРЗД). Голяма част от интервюираните съобщават, че е извършена оценка на въздействието върху защитата на данните съгласно изискванията на законодателството. Същевременно тези оценки са изпълнени под различни форми. Освен това предишните оценки, когато са извършвани такива, се съсредоточават предимно върху техническите аспекти. Те рядко разглеждат потенциалните въздействия върху основните права. Според някои от интервюираните не се извършват оценки на въздействието върху основните права, когато дадена система с ИИ не засяга или не изглежда да засяга неблагоприятно основните права.

Проучването показва, че познанията на интервюираните за основните права, ако се изключат защитата на данните и, до известна степен, недопускането на дискриминация, са ограничени. Същевременно повечето признават, че използването на ИИ оказва въздействие върху основните права. Някои от интервюираните посочват, че техните системи не засягат основните права, което до известна степен е свързано със задачите, за които се използват системите с ИИ.

Всички респонденти са запознати с въпросите, свързани със защитата на личните данни. Повечето също така осъзнават, че при използването на ИИ в общия случай дискриминацията би могла да бъде проблем. Същевременно точният смисъл и приложимостта на правата, свързани със защитата на личните данни и недопускането на дискриминация, остават неясни за голяма част от респондентите.

Констатациите от проучването показват различия между частния и публичния сектор. Интервюираните от частния сектор често са по-малко осведомени за по-широкия спектър от основни права, които биха могли да бъдат засегнати. Проблемите, свързани със защитата на личните данни, са добре известни на частния сектор. Останалите права, например правата, свързани с недопускане на дискриминация или достъп до правосъдие, са по-малко известни сред представителите на предприятия, работещи с ИИ. Някои са напълно запознати с потенциалните проблеми. Други казват, че техните клиенти са тези, които носят отговорността да проверяват за наличие на проблеми при спазването на основните права.



СТАНОВИЩЕ 3 НА FRA

ЕС и държавите членки трябва да осигурят въвеждането на ефективни системи за отчетност, за да наблюдават и, ако е необходимо, ефективно да преодоляват всякакво отрицателно въздействие на системите с ИИ върху основните права. В допълнение към оценките на въздействието върху основните права (вж. становище 2 на FRA) те трябва да разгледат и възможността за въвеждане на конкретни мерки за защита, за да се гарантира ефективността на режима за отчетност. Това може да включва правно изискване за предоставяне на достатъчно информация, която да позволи оценка на въздействието на системите с ИИ върху основните права. По този начин компетентните органи биха могли да осъществяват ефективно наблюдение и надзор върху спазването на правата на човека.

Освен това ЕС и държавите членки трябва се възползват по-пълноценно от съществуващите експертни структури за надзор за защита на основните права при използване на ИИ. Те включват органите за защита на данните, органите по въпросите на равенството, националните институции за правата на човека, институциите на омбудсманите и органите за защита на потребителите.

Трябва да бъдат предвидени допълнителни ресурси за създаване на ефективни системи за отчетност чрез повишаване на квалификацията и разнообразяване на персонала, работещ за надзорните органи. Това ще им даде възможност да се справят с комплексни въпроси, свързани с разработването и използването на ИИ.

По същия начин съответните органи трябва да разполагат с достатъчно ресурси, правомощия и най-вече квалификация, за да предотвратяват и извършват оценка на нарушенията на основните права и за да подкрепят ефективно тези, чиито основни права са засегнати от ИИ.

Улесняването на сътрудничеството между съответните органи на национално и европейско равнище може да помогне за обмяна на знания и опит. От помощ би могло да бъде и ангажирането на други действащи лица със съответните познания, например специализирани организации на гражданското общество. При изпълнение на такива действия на национално ниво държавите членки трябва да разгледат възможността да използват наличните механизми за финансиране от Съюза.

Осигуряване на ефективен надзор и цялостна отчетност

Предприятията и държавните администрации, които разработват и използват ИИ, осъществяват контакт с различни органи, отговарящи за надзора върху системите, свързани с ИИ, в рамките на съответните си задължения и сектори. Тези органи включват органите за защита на данните. Същевременно ползващите ИИ невинаги са сигурни кои са органите, отговарящи за надзора върху системите с ИИ.

В съответствие с добре установените международни стандарти в областта на правата на човека, например член 1 от Европейската конвенция за правата на човека (ЕКПЧ) и член 51 от Хартата, държавите са задължени да гарантират правата и свободите на хората. За да изпълнят ефективно това изискване, държавите трябва, освен всичко друго, да въведат ефективни механизми за наблюдение и изпълнение. Това се отнася в еднаква степен и за ИИ.

По отношение на наблюдението констатациите сочат важната роля на специализираните органи, създадени в конкретните сектори, които отговарят и за надзора върху ИИ в рамките на своите задължения. Това включва например надзора в областта на банковото дело или органите за защита на данните. Съществуват различни такива органи, които са потенциално свързани с надзора върху ИИ от гледна точка на основните права. От друга страна, отговорностите на органите по отношение на надзора върху ИИ остават неясни за голяма част от интервюираните от частния и държавния сектор.

Използването на ИИ от държавните администрации понякога е обект на проверка като част от редовните одити. Частните дружества в някои конкретни сектори също разполагат със специализирани органи за надзор, например в областта на здравеопазването или финансовите услуги. Освен това тези органи проверяват използването на ИИ и свързаните с него технологии, например като част от схемите си за сертифициране. Интервюираните от частния сектор изразиха желание за създаването на органи, които ще могат да дават експертни съвети относно възможностите и законността на потенциалните възможности за използване на ИИ.

В ЕС има добре развити независими органи, чието задължение е да защитават и насърчават основните права. Те включват органите за защита на данните, органите по въпросите на равенството, националните институции за правата на човека и институциите на омбудсманите. Проучването показва, че тези, които използват или планират да използват ИИ, често са се свързвали с различни органи във връзка с използването на ИИ, например органи за защита на потребителите.

Най-често ползвателите на ИИ са се свързвали с органи за защита на данните за насоки, помощ и одобрение в случаите, когато е налице обработване на лични данни. Интервюираните експерти подчертават значението на органите за защита на данните с цел надзор върху системите с ИИ във връзка с използването на лични данни. Същевременно те отбелязват, че органите за защита на данните не разполагат с достатъчно ресурси за тази задача и им липсват конкретните експертни познания по въпросите, свързани с ИИ.

Експертите, включително тези, които работят за надзорни органи, например органите по въпросите на равенството и органите за защита на данните, са на мнение, че е необходимо повишаване на квалификацията на съществуващите надзорни органи, за да могат да осъществяват ефективен надзор по въпросите, свързани с ИИ. Според експертите това може да се окаже предизвикателство предвид факта, че ресурсите на тези органи вече са доста ограничени. Те подчертаха и важната роля на съответните организации на гражданското общество, специализирани в областта на технологиите, цифровите права и алгоритмите. Тези организации биха могли да подобрят отчетността при използване на системи с ИИ.



НЕДОПУСКАНЕ НА ДИСКРИМИНАЦИЯ, ЗАЩИТА НА ЛИЧНИТЕ ДАННИ И ДОСТЪП ДО ПРАВОСЪДИЕ: ТРИ ХОРИЗОНТАЛНИ ТЕМИ

Проучването показва, че използването на ИИ засяга различни основни права. Освен свързаните с контекста специфични аспекти, които засягат различни права в различна степен, темите относно основните права, които възникнаха по време на проучването и които многократно се отнасят за по-голямата част от случаите с ИИ, включват: необходимостта да се осигури използване на ИИ без допускане на дискриминация (право на недискриминация), изискването за законно обработване на данните (право на защита на личните данни) и възможността за подаване на оплакване за решения, основани на ИИ, и търсене на правна защита (право на ефективни правни средства за защита и на справедлив съдебен процес).

Двете най-важни основни права, подчертани по време на интервютата, са защитата на личните данни и недопускането на дискриминация. Освен това неколкотократно се засяга темата за ефективните начини за подаване на оплакване относно използването на ИИ, свързана с правото на справедлив съдебен процес и ефективни правни средства за защита. Следващите три становища на FRA, които отразяват тези констатации, следва да бъдат разглеждани успоредно с останалите становища, които призовават към по-всеобхватно отчитане и реагиране спрямо пълния спектър от основни права, засегнати от ИИ.

СТАНОВИЩЕ 4 НА FRA

Държавите — членки на ЕС, трябва да разгледат възможността да насърчат дружествата и държавната администрация да извършват оценка на потенциално дискриминационните резултати при използване на системи с ИИ.

Европейската комисия и държавите членки трябва да обмислят предоставянето на финансиране за целенасочени изследвания върху потенциално дискриминационните въздействия от използването на ИИ и алгоритми. Тези изследвания биха могли да се възползват от адаптирането на установени изследователски методологии от социалните науки, които се използват за установяване на потенциална дискриминация в различни области, вариращи от подбор на персонал до създаване на профили на клиенти.

Въз основа на резултатите от тези изследвания, трябва да се разработят насоки и инструменти, които да помогнат на тези, които използват ИИ, да установяват евентуални дискриминационни резултати.

Специфични мерки за защита за осигуряване на недопускане на дискриминация при използване на ИИ

Интервюираните рядко споменават извършването на подробни оценки за евентуална дискриминация при използване на ИИ. Това предполага липса на задълбочени оценки за наличието на такава дискриминация при автоматизираното вземане на решения.

Задължението за зачитане на принципа на недопускане на дискриминация е заложено в член 2 от Договора за ЕС, член 10 от Договора за функционирането на ЕС (който изисква от Съюза да се бори срещу дискриминацията на различна основа) и член 20 и 21 от Хартата (равенство пред закона и недопускане на дискриминация на различна основа). Някои по-конкретни и подробни разпоредби в няколко директиви на ЕС също съдържат този принцип с различен обхват на приложение.

Автоматизацията и използването на ИИ може до голяма степен да увеличи ефективността на услугите и да увеличи мащаба на задачи, които хората

не биха били в състояние да изпълнят. Същевременно е необходимо да се гарантира, че услугите и решенията, основани на ИИ, не са дискриминационни. Отчитайки това, Европейската комисия наскоро подчерта необходимостта от допълнително законодателство за спазване на принципа за недопускане на дискриминация при използване на ИИ в **Плана за действие на ЕС за борба с расизма за периода 2020—2025 г.**

Повечето интервюирани като цяло са наясно, че е възможна дискриминация. Въпреки това те самите рядко са повдигали този въпрос. Само някои от тях смятат, че системите им всъщност биха могли да допуснат дискриминация.

Освен това интервюираните рядко споменават за подробни оценки за потенциална дискриминация, което означава, че липсва задълбочена оценка за потенциалната дискриминация.

Общоприето мнение е, че пропускането на информация за някои данни, които са обект на защита, например пол, възраст или етнически произход, може да гарантира, че системата с ИИ няма да допуска дискриминация. Това обаче не е непременно вярно. Информация, потенциално посочваща характеристики, които са обект на защита (показатели), която често може да бъде открита в набори от данни, би могла да доведе до дискриминация.

В някои случаи системите с ИИ могат да се използват също за изпитване и установяване на дискриминационно поведение, което може да бъде кодирано в наборите от данни. Същевременно много малка част от интервюираните споменават възможността за събиране на такава информация за групи в неравностойно положение за установяване на потенциална дискриминация. При липсата на задълбочен анализ за потенциалната дискриминация при действителното използване на системи с ИИ почти липсват също обсъждане и анализ на потенциалния положителен ефект от използването на алгоритми, които да направят решенията по-справедливи. Освен това нито един от интервюираните, работещи върху ИИ, не споменава използването на ИИ за установяване на евентуална дискриминация като положителен резултат с оглед на това, че дискриминацията би могла да се установи по-лесно при анализиране на данните за възможни отклонения.

Тъй като установяването на потенциална дискриминация чрез използване на ИИ и алгоритми си остава предизвикателство, а интервюираните не отделят голямо внимание на този въпрос, са необходими различни мерки за преодоляване на този проблем. Те включват изискването да се разгледат въпросите, свързани с дискриминацията, при оценката на използването на ИИ, и да се инвестира в допълнителни проучвания за потенциална дискриминация, които да използват разнообразен набор от методологии.

Те биха могли да включват например изпитване за дискриминация. По този начин може да се надгражда върху подобни установени методологии за изпитване за отклонения в ежедневието, например тези, свързани с кандидатстване за работа, при които името на кандидата се променя така, че (косвено) да показва етническия произход. По отношение на приложенията с ИИ тези изпитвания биха могли да включват евентуално създаване на фалшиви профили за онлайн инструменти, които да се различават само по отношение на данните, които са обект на защита. По този начин резултатите биха могли да бъдат проверени с оглед на възможността за дискриминация. От полза за изследванията би могъл да бъде и задълбочен статистически анализ за установяване на различията в наборите от данни, свързани със защитени групи, който би могъл да се използва като основа за изследване на евентуалната дискриминация.

На последно място, някои интервюта по проучването подчертават, че резултатите от комплексни алгоритми за машинно самообучение често са много трудни за разбиране и обяснение. Следователно допълнителните изследвания с цел по-ясно разбиране и обясняване на тези резултати (т.нар. „обясним ИИ“) също биха могли да помогнат за подобряване на установяването на дискриминация при използване на ИИ.

```
mod.use_x = False
mod.use_y = False
mod.use_z = True
objects.active = modifier of
function at the end -add back the def
ob.select= 1
ob.select=1
EXPORTSYMBOL(groupsFree);
/* export the groupinfo to a user-space array */
EXPORTSYMBOL(groupsFree);
int groups_touser(gid_t_user *grouplist,
/* export the groupinfo to a user-space array */
const struct group_info *group_info)
static int groups_touser(gid_t_user *grouplist,
const struct group_info *group_info)
{
int i;
```

СТАНОВИЩЕ 5 НА FRA

Европейският комитет по защита на данните (ЕКЗД) и Европейският надзорен орган по защита на данните (ЕНОЗД) трябва да разгледат възможността да предоставят допълнителни насоки и подкрепа за ефективно прилагане на разпоредбите на ОРЗД, които се отнасят пряко за използването на ИИ за защита на основните права, по-специално във връзка със значението на личните данни и използването им в ИИ, включително в наборите от данни за обучение на ИИ.

Съществува висока степен на несигурност относно значението на автоматизираното вземане на решения и правото на човешка преценка, свързани с използването на ИИ и автоматизираното вземане на решения. Следователно ЕКЗД и ЕНОЗД трябва да обърнат внимание и на допълнителното поясняване на понятията за „автоматизирано вземане на решения“ и „човешка преценка“ в случаите им на споменаване в правото на Съюза.

В допълнение, националните органи за защита на данните трябва да предоставят практически насоки относно относимостта на разпоредбите за защита на личните данни към използването на ИИ. Тези насоки биха могли да включват препоръки и контролни списъци въз основа на конкретни примери за използване на ИИ, за да подпомогнат спазването на разпоредбите за защита на личните данни.

Още насоки относно защитата на данните

Необходима е повече яснота за обхвата и значението на законовите разпоредби относно автоматизираното вземане на решения.

Защитата на личните данни е от решаващо значение при разработването и използването на ИИ. Член 8, параграф 1 от Хартата и член 16, параграф 1 от Договора за функционирането на ЕС гласят, че всеки има право на защита на личните си данни. ОРЗД и Директивата относно правоприлагането (Директива (ЕС) 2016/680) допълнително разглеждат това право и включват множество разпоредби, приложими за използването на ИИ.

Интервюираните посочват, че повечето прилагани от тях системи с ИИ използват лични данни, което означава, че защитата на личните данни е засегната по различни начини. Същевременно някои приложения, според интервюираните, не използват лични данни или използват само анонимизирани данни и съответно законодателството за защита на личните данни не би се отнасяло за тях. Ако се използват лични данни, се прилагат всички принципи и разпоредби, свързани със защита на личните данни.

Този доклад набелязва една важна тема, свързана със защитата на личните данни, която е относима и към други основни права, свързани с автоматизираното вземане на решения. Според проучване на Евробарометър само 40 % от европейците знаят, че имат право на мнение в случаите, когато решенията са автоматизирани. Осведомеността за това право

е значително по-висока сред тези, които работят с ИИ — повечето от интервюираните лица повдигат този въпрос. Същевременно голяма част от тях, включително експерти, твърдят, че е необходима повече яснота за обхвата и значението на законовите разпоредби относно автоматизираното вземане на решения.

В областта на социалните помощи интервюиращите споменават само един пример за напълно автоматизирани решения на основата на правило. Всички останали приложения, които споменават, са обект на човешка преценка. Интервюираните от държавната администрация подчертават важността на човешката преценка при всички решения. Същевременно те рядко описват какво всъщност включва тази човешка преценка и по какъв начин е използвана останалата информация при прегледа на резултатите от системите с ИИ.

Въпреки че интервюираните са на различно мнение по въпроса дали съществуващото законодателство е достатъчно, голяма част от тях призовават към по-конкретно тълкуване на настоящите правила за защита на личните данни по отношение на автоматизираното вземане на решения, както е заложено в член 22 от ОРЗД.

Ефективен достъп до правосъдие в случаи, свързани с решения, основани на ИИ

За да оспорват ефективно решения, основани на използването на ИИ, е необходимо хората да знаят за използването на ИИ, както и начина и мястото, където да подадат оплакване. Организациите, които използват ИИ, трябва да могат да обяснят принципа на своите системи с ИИ и решенията, основани на ИИ.

Достъпът до правосъдие е едновременно процес и цел и е от решаващо значение за лицата, които желаят да се възползват от други процесуални и материални права. Той е свързан с някои основни права на човека. Те включват правото на справедлив съдебен процес и на ефективни правни средства за защита съгласно членове 6 и 13 от Европейската конвенция за правата на човека и член 47 от Хартата на основните права на ЕС. Следователно понятието за достъп до правосъдие задължава държавите да гарантират правото на всяко лице да се яви пред съда или, при някои обстоятелства, пред орган за алтернативно разрешаване на спорове, за да се възползва от средствата за правна защита, ако бъде установено, че правата на това лице са нарушени.

Според тези стандарти лице, което е жертва на нарушаване на правата на човека в резултат от разработването или използването на система с ИИ от държавен или частен субект, трябва да разполага с достъп до средства за правна защита пред национален орган. В съответствие с приложимата съдебна практика съгласно член 47 от Хартата и член 13 от Европейската конвенция за правата на човека, средството за правна защита трябва да бъде „ефективно както от практическа, така и от правна гледна точка“.

Констатациите от проучването установяват следните предпоставки, за да бъде средството за правна защита ефективно от практическа гледна точка в случаите, свързани със системи с ИИ, и тяхното въздействие върху основните права: когато се използва ИИ, всеки трябва да е запознат с това и да бъде информиран как и къде да подаде оплакване. Организациите, които използват ИИ, трябва да осигурят информиране на обществеността за своята система с ИИ и за решенията, основани на нея.

Констатациите показват, че разясняването на системите с ИИ и начина, по който вземат решения, на разбираем език може да представлява предизвикателство. Правата на интелектуална собственост може да възпрепятстват предоставянето на подробна информация за начина на работа на даден алгоритъм. Освен това някои системи с ИИ са комплексни. Това затруднява предоставянето на смислена информация за начина на работа на системата и съответните решения.

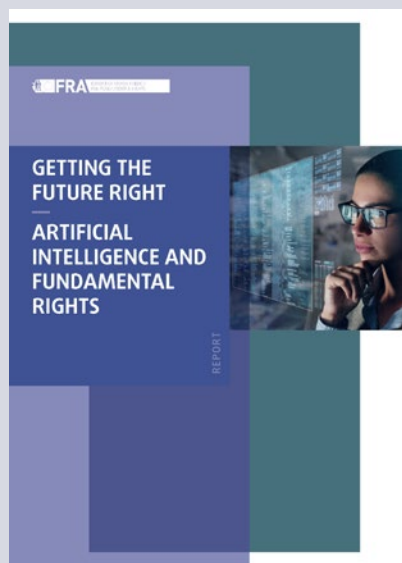
СТАНОВИЩЕ 6 НА FRA

Законодателят на Съюза и държавите членки трябва да осигурят ефективен достъп до правосъдие за лицата в случаите, свързани с решения, основани на ИИ.

За да се гарантира, че наличните средства за правна защита са достъпни на практика, законодателят на Съюза и държавите членки трябва да обмислят въвеждането на законово задължение за държавната администрация и частните дружества, които използват системи с ИИ, да предоставят информация за работата на своите системи с ИИ на тези, които търсят правна защита. Това включва информация за начина, по който системите с ИИ достигат до автоматизираните решения. Това задължение ще помогне за постигане на равни процесуални възможности за лицата, търсещи справедливост. Освен това то би подкрепило ефективността на външното наблюдение и надзор върху спазването на правата на човека от системите с ИИ (вж. становище 3 на FRA).

С оглед на трудността за разясняване на комплексните системи с ИИ Съюзът, съвместно с държавите членки, трябва да обмисли разработването на насоки, подкрепящи усилията за прозрачност в тази област. В изпълнение на това задължение те трябва да изхождат от познанията на националните органи за правата на човека и организациите на гражданското общество, които са активни в тази област.

За да се справят с този проблем, някои от интервюираните дружества изцяло избягват използването на комплексни методи за вземането на някои решения поради това, че не биха могли да обяснят решенията. Като алтернатива те използват по-прости методи за анализ на данни за същия проблем, за да получат известна информация за основните фактори, които влияят върху определени резултати. Някои от интервюираните представители на частния сектор акцентират върху положените усилия за постепенно подобряване на тяхното разбиране за технологиите с ИИ.



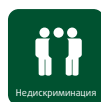
Настоящото резюме представя основните констатации от доклада на FRA със заглавие „Правилно изграждане на бъдещето – изкуствен интелект и основни права“. Основният доклад е достъпен на [уебсайта на FRA](#).



ПОПУЛЯРИЗИРАНЕ И ЗАЩИТА НА ВАШИТЕ ОСНОВНИ ПРАВА В ЕС

Докладът на FRA относно изкуствения интелект и основните права представя конкретни примери за начина, по който дружествата и държавните администрации в ЕС използват или се опитват да използват ИИ. В него са разгледани четири основни области: социални помощи, прогнозиране в полицейската област, здравни услуги и целево рекламиране. Обсъждат се потенциалните отражения върху основните права и се анализира по какъв начин са отчетени тези права при използването или разработването на приложения с ИИ.

В настоящото резюме се представят основните заключения от доклада. Те могат да послужат за информация при създаване на политики на равнището на ЕС и на национално равнище за регламентиране на използването на инструменти с ИИ в съответствие с правата на човека и основните права.



FRA – АГЕНЦИЯ НА ЕВРОПЕЙСКИЯ СЪЮЗ ЗА ОСНОВНИТЕ ПРАВА
Schwarzenbergplatz 11 – 1040 Виена – Австрия
Тел. +43 158030-0 – Факс +43 158030-699

fra.europa.eu

[facebook.com/fundamentalrights](https://www.facebook.com/fundamentalrights)
twitter.com/EURightsAgency
[linkedin.com/company/eu-fundamental-rights-agency](https://www.linkedin.com/company/eu-fundamental-rights-agency)



Служба за публикации
на Европейския съюз