

**Регулиране на изкуствения интелект в
общественото здравеопазване с цел
предотвратяване на системни
пристрастия, основани на пол, цвят на
кожата или етнически произход**

**Регулаторен модел, съобразен с Хартата
на основните права на Европейския съюз**

Полиси документ



Съфинансирано от
Европейския съюз

Въведение

Изкуственият интелект (ИИ) все по-широко се интегрира в здравните системи в Европа, като подпомага диагностиката, вземането на клинични решения, планирането на лечението и прогнозирането на риска за пациентите. Тези технологии предлагат значителни възможности за повишаване на ефективността на здравеопазването и подобряване на здравните резултати. Същевременно има все повече данни, че **системите с ИИ могат да възпроизвеждат или задълбочават съществуващите неравенства в здравеопазването**.¹

Съществуват няколко механизма, чрез които могат да възникнат системни пристрастия при вземането на решения в здравеопазването, подпомагани от ИИ. Сред тях са дисбаланси в обучаващите данни, грешки при измерването на клинични променливи, използването на неподходящи прокси индикатори и различията между средата, в която системите са разработени, и условията на тяхното реално внедряване и използване.² Системни пристрастия могат да възникнат и когато системите с ИИ се въвеждат в клиничните процеси без адекватно обучение на здравните специалисти или без ефективни механизми за човешки надзор и отчетност, които да проследяват тяхното функциониране сред разнообразни групи пациенти.³

Освен това емпиричните изследвания показват, че подобни **системни пристрастия могат да имат сериозни последици в клиничната практика**. Проучванията отчитат трайни неравенства при диагностицирането и лечението на редица значими заболявания. Жените и пациентите от етнически малцинства със сърдечносъдови заболявания често получават по-късна или по-малко интензивна диагностика и лечение в сравнение с мъжете и пациентите от мнозинствените групи.⁴ Различията,

¹ Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y., *Изкуствен интелект в здравеопазването: минало, настояще и бъдеще*, *Stroke and Vascular Neurology* 2(4): 230–243, 2017.

² Vokinger, K. N., Feuerriegel, S., & Kesselheim, A. S., *Смекчаване на системни пристрастия в машинното обучение за медицина*, *Communications Medicine* 1, 25, 2021.

³ Panesar, S., Klot, M., Parrish, R., Fernandez-Miranda, J., Cagle, Y., & Britz, G.W., *Обещанията и рисковете на изкуствения интелект в неврохирургията*, *Neurosurgery* 87(1): 33–44, 2020.

⁴ Balla, S., Gomez, S.E., & Rodriguez, F., *Неравенства в сърдечно-съдовите грижи и резултати при жени от расови/етнически малцинствени групи*, *Current Treatment Options in Cardiovascular Medicine* 22(12): 75, 2020.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI

основани на цвят на кожата или етнос, оказват влияние върху грижите и резултатите при диабет.⁵ Психичните разстройства, като депресията, са недостатъчно диагностицирани и недостатъчно лекувани както при жените, така и сред малцинствените групи от населението.⁶ Показателен пример представлява проучване, публикувано в *Journal of Medical Internet Research*, което установява различия в ефективността на алгоритмите за машинно обучение, използвани за прогнозиране на сърдечни заболявания, в зависимост от пола на пациентите. По-конкретно, жените са били системно недостатъчно представени в обучаващите данни, а моделите са показвали по-висок дял на фалшиво отрицателни резултати при жените пациенти. Това демонстрира как пристрастията в данните могат да доведат до **несправедливи диагностични резултати**.⁷ В допълнение към това, друго изследване, насочено към равенството между хората с различен цвят на кожата, установява, че модели на ИИ, обучени да прогнозират тежестта на депресията въз основа на езика, използван в социалните мрежи, постигат добри резултати при бели участници, но значително по-слаби резултати при чернокожи участници. Това разкрива различия в ефективността на моделите по линия на цвят на кожата, които отразяват особености в езиковото изразяване, останали извън обхвата на използваните модели.⁸

Преодоляването на риска от подобни системни пристрастия изисква **регулаторни модели, които превеждат международните и европейските етични и правни принципи в конкретни механизми за управление**, приложими на практика от здравните институции. В рамките на Европейския съюз (ЕС) регулаторната рамка се развива динамично с приемането на Акта за изкуствения интелект (AI Act), който въвежда задължения за високорисковите системи с ИИ и за организациите, които ги внедряват и използват, включително болниците.⁹ Въпреки това превръщането на тези правни изисквания в практически процедури и

⁵ Young, C. & Myers, A. K., *Расови и етнически неравенства в клиничните грижи и управлението на диабет: наративен преглед*, *Endocrine Practice* 29(4): 295–300, 2023.

⁶ Sorkin, D.H., Ngo-Metzger, Q., Billimek, J., August, K.J., Greenfield, S., & Kaplan, S.H., *Недостатъчно диагностицирана и недостатъчно лекувана депресия сред пациенти с диабет тип 2 от различни расови и етнически групи*, *Diabetes Care* 34(3): 598–600, 2011.

⁷ Straw, I., Rees, G., & Nachev, P., *Различия в представянето, основани на биологичен и социален пол, в алгоритмите за машинно обучение за предсказване на сърдечни заболявания: проучвателно изследване*, *Journal of Medical Internet Research* 26: e46936, 2024.

⁸ Rai, S., Stadel, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C., *Ключови езикови маркери на депресията в социалните медии зависят от расата*, *Proceedings of the National Academy of Sciences* 121(14): e2319837121, 2024.

⁹ European Union, *Регламент (ЕС) 2024/1689 на Европейския парламент и на Съвета за установяване на хармонизирани правила относно изкуствения интелект (Акт за изкуствения интелект)*, PE/24/2024/REV/1, *Official Journal of the European Union*, 13 June 2024.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



организационни механизми в рамките на здравните институции представлява съществено предизвикателство.

Настоящият документ предлага **регулаторен модел за управление на биомедицинския ИИ, който е съобразен с достиженията на правото на ЕС**, включително с Хартата на основните права на ЕС, и е предназначен за приложение в системите на общественото здравеопазване и гражданските организации. Предложеният модел се основава на европейските регулаторни инструменти, включително Акта за ИИ, Регламента за медицинските изделия (MDR), Регламента за медицинските изделия за ин витро диагностика (IVDR), Общия регламент относно защитата на данните (GDPR) и нововъзникващото Европейско пространство на здравни данни (EHDS). Той превръща принципите на основните права, и по-специално човешкото достойнство, равенството, недискриминацията и правото на здравеопазване, в практически гаранции, приложими през целия жизнен цикъл на системите с ИИ. Позоваванията на избрани примери от държави извън ЕС са включени единствено когато предоставят приложими добри практики и полезни изводи, които могат да бъдат адаптирани в европейски контекст. Тези примери имат допълващ характер спрямо европейската правна рамка и подхода, основан на защитата на основните права, върху които е изграден предложеният модел.

Освен това регулаторният модел се основава на утвърдени подходи за регулиране и управление на риска, използвани в други високорискови сектори, включително управлението на безопасността в авиацията, управлението на риска при финансовите модели, фармакологичната бдителност в регулацията на лекарствените продукти и оценките на въздействието на алгоритмите в публичната администрация. Тези подходи предоставят ценни насоки за изграждането на рамки за управление, способни да адресират сложните рискове, възникващи през целия жизнен цикъл на биомедицинския ИИ. В настоящия модел те са адаптирани така, че изрично да интегрират съображенията за равенство между половете и между хората с различен цвят на кожата във всички етапи от жизнения цикъл на системите с ИИ.

Моделът предоставя на здравните институции и гражданските организации **практическа рамка за управление, обхващаща възлагането и закупуването, валидирането, внедряването, наблюдението и механизмите за отчетност** на биомедицинските системи с ИИ. Чрез интегрирането на гаранции за защита на основните права в оперативните процеси тази рамка има за цел да намали риска от системни пристрастия, основани на пол и цвят на кожата или етнически произход, при вземането на

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



решения в здравеопазването, подпомагани от ИИ, и да допринесе за по-справедливи и равни здравни резултати в рамките на ЕС. В допълнение документът **предлага набор от практически политики, инструменти и образци на документи, които болниците и гражданските организации могат да адаптират към различни национални контексти с цел последователно прилагане на модела в реални условия.**

Европейската правна рамка за ИИ в здравеопазването

В рамките на ЕС има изградена многостепенна регулаторна среда за биомедицинския ИИ, състояща се от четири взаимодопълващи се правни слоя: Акта за ИИ, законодателството относно медицинските изделия и свързаните с него насоки, правилата за защита на данните и рамката за управление на здравните данни. Акта за ИИ въвежда регулаторна рамка, основана на риска, за системите с ИИ. Много приложения в здравеопазването, включително инструментите за подпомагане на клиничните решения и диагностичните алгоритми, попадат в категорията на високорисковите системи с ИИ и следва да отговарят на строги изисквания, свързани с управлението на риска, управлението на данните, прозрачността, човешкия надзор и точността. Сред приложимите изисквания са също идентифицирането и ограничаването на риска от дискриминационни резултати чрез сравнителен анализ на различни групи от населението, като биологичният и социалният пол, етническият произход и други социални категории се разглеждат като хоризонтални променливи. **Здравните институции, които внедряват и използват системи с ИИ, също имат специфични задължения съгласно Акта за ИИ.** Те включват осигуряване на използването на системите в съответствие с инструкциите на доставчика, наблюдение на тяхното функциониране, поддържане на оперативни дневници и докладване на сериозни инциденти на компетентните органи, включително когато такива инциденти могат да засегнат непропорционално определени групи пациенти.

Що се отнася до прилагането на Акта за ИИ, още през 2025 г. започнаха да се прилагат правилата относно забранените практики, както и част от изискванията, свързани с надзора и управлението.¹⁰ Следващият основен етап настъпва през август 2026 г., когато по-голямата част от задълженията по регламента ще станат приложими, включително тези, свързани с високорисковите системи по Приложение III и задълженията за прозрачност

¹⁰ European Commission, [Хронология за прилагането на Акта на ЕС за изкуствения интелект](#), Brussels: European Commission, 2024.



по член 50. За високорисковите системи с ИИ, интегрирани в регулирани продукти – категория, в която попадат много от системите с ИИ, използвани в медицинските изделия – прилагането на съответните изисквания е предвидено да започне през август 2027 г.

За болниците най-важните от практическа гледна точка разпоредби на Акта за ИИ са:

- **Изисквания към високорисковите системи.** Доставчиците следва да внедрят система за управление на риска, обхващаща целия жизнен цикъл на системата (чл. 9), както и практики за управление на данните, включващи изрични механизми за откриване и ограничаване на системни пристрастия (чл. 10). Това включва оценка на представителността на данните и на потенциалните различия в ефективността на системите спрямо различни групи от населението чрез анализ по биологичен и социален пол, както и цвят на кожата или етнически произход. Изискванията включват също мерки за прозрачност и предоставяне на информация на организациите, които внедряват системите (чл. 13), механизми за човешки надзор, предназначени да предотвратяват сляпото доверие в автоматизирани решения и да позволяват намеса, спиране или отменяне на решенията на системата (чл. 14), както и мерки за гарантиране на точността, устойчивостта и киберсигурността на системите (чл. 15). Последните следва изрично да отчитат риска от цикли на обратна връзка и специфични за ИИ атаки, които биха могли да задълбочат съществуващите неравенства, ако не бъдат адекватно адресирани.
- **Задължения на разпоредителите (deployer obligations).** Болниците и другите разпоредители трябва да използват високорискови системи с ИИ в съответствие с инструкциите на доставчика (чл. 26); да определят компетентен човешки надзор (чл. 26); да гарантират, че входните данни са представителни и пълни, когато разпоредителят контролира въвеждането на данни (чл. 26), включително (доколкото е възможно) отчитане на демографското разнообразие сред пациентите; да наблюдават функционирането на системата (чл. 26); да съхраняват автоматично създадените записи за дейността на системата за срок от най-малко шест месеца (чл. 26); и своевременно да докладват сериозни инциденти на доставчиците и компетентните органи (чл. 26, 73), особено когато съществува риск от несправедливи клинични резултати.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



- **Оценка на въздействието върху основните права** (Fundamental Rights Impact Assessment – FRIA). Преди внедряването на определени високорискови системи с ИИ, органи, регулирани от публичното право (както и частни организации, предоставящи обществени услуги), са длъжни да извършат такава оценка, като опишат процесите, засегнатите групи, рисковете, мерките за човешки надзор и механизмите за смекчаване на рисковете или за подаване на жалби, и след това да уведомят компетентния орган (чл. 28). Това изискване е пряко относимо към публичните болници в много държави членки, тъй като предоставя механизъм за идентифициране и адресиране, преди внедряването, на рискове от дискриминация и неравноправни въздействия, оценени чрез анализи по биологичен и социален пол, както и по цвят на кожата или етнически произход.

Освен Акта за ИИ, **системите с ИИ, използвани в медицински контекст, подлежат и на законодателството за медицинските изделия, както и на свързаните насоки**, включително Регламент (ЕС) за медицинските изделия (Medical Device Regulation – MDR)¹¹ и Регламент (ЕС) за ин витро диагностичните медицински изделия (In Vitro Diagnostic Regulation – IVDR),¹² които уреждат безопасността и ефективността на медицинските технологии. Тези регламенти изискват клинична оценка и надзор след пускане на пазара за софтуер като медицинско изделие, включително инструменти, базирани на ИИ, като се обръща внимание на справедливата ефективност сред различни пациенти, доколкото наличните данни позволяват. Освен това Насоките на Групата за координация по медицинските изделия (Medical Device Coordination Group – MDCG) относно квалификацията и класификацията на софтуер съгласно MDR/IVDR предоставят указания как софтуерът се квалифицира и класифицира по тези регламенти.¹³ В тях се подчертава, че класификацията зависи от предвидената употреба, а не от местоположението на софтуера (например облачно базиран или вграден), и се включват примери, извлечени от болнични информационни системи и софтуер за клинична подкрепа при вземане на решения, като се акцентира върху необходимостта да се отчитат въздействията върху пациенти с различни демографски характеристики. Насоките на на Групата за

¹¹ European Union, *Regulation (EU) 2017/745 on Medical Devices (MDR)*, Official Journal of the European Union, L117, Brussels: EU, 2017.

¹² European Union, *Регламент (ЕС) 2017/746 относно медицинските изделия за ин витро диагностика (IVDR)*, Official Journal of the European Union, L117, Brussels: EU, 2017.

¹³ Medical Device Coordination Group, *Насоки относно квалификацията и класификацията на софтуер съгласно Регламент (ЕС) 2017/745 (MDR) и Регламент (ЕС) 2017/746 (IVDR)*, Brussels: MDCG, 2019.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



координация по медицинските изделия относно клиничната и експлоатационната оценка на софтуер за медицински изделия предоставят допълнителна подкрепа, като утвърждават очакването доказателствата да бъдат съразмерни с клиничната роля на софтуера, включително доказателства, че ефективността на ИИ е последователна сред пациенти с различен биологичен и социален пол или цвят на кожата и етнически произход.¹⁴

Връзката между Акта за ИИ и съществуващото право за медицинските изделия е изяснена в Насоките относно взаимодействието между Акта за ИИ и MDR/IVDR.¹⁵ Документът обяснява, че Акта за ИИ се прилага паралелно с тези два регламента. На практика това означава, че много **медицински системи с ИИ трябва да отговарят едновременно на двете регулаторни рамки**. Насоките също така изясняват кога медицинският ИИ се квалифицира като високорисков по смисъла на Акта за ИИ, което обикновено е налице, когато самата ИИ система представлява медицинско изделие (или негов компонент, свързан с безопасността) и следователно подлежи на оценяване на съответствието от трета страна (от компетентен орган). По отношение на обществените поръчки в болничната практика това има практически последици. Болниците следва да изискват от доставчиците да доказват регулаторния статус на своите системи, включително класификация по MDR/IVDR, участие на компетентен орган и съответствие с изискванията за висок риск по Акта за ИИ, както и документация за представянето на системата при различни групи от населението и мерки за смекчаване на пристрастията. Договорите за обществени поръчки следва също така да изискват от доставчиците да предоставят необходимата документация, информация за прозрачност и механизми за надзор, изисквани от разпоредителите по Акта за ИИ, включително разпоредби, гарантиращи справедливи резултати за всички пациенти.

Няколко **допълнителни правни инструмента на ЕС също са от значение за управлението на ИИ в здравеопазването на практика**. Общият регламент относно защитата на данните остава основно ограничение за биомедицинския ИИ, тъй като здравните данни представляват специална категория лични данни, а внедряванията в болнична среда често изискват оценки на въздействието върху защитата на данните (Data Protection Impact Assessments – DPIA). Актът за ИИ (чл. 26, пар. 9 и чл. 27, пар. 4) изрично свързва практиките на разпоредителите с изискванията за оценка на

¹⁴ Medical Device Coordination Group, *Насоки относно клиничната оценка (MDR) и оценката на представянето (IVDR) на софтуер в медицински изделия*, Brussels: MDCG, 2020.

¹⁵ Medical Device Coordination Group & European AI Board, *Насоки относно взаимодействието между MDR/IVDR и Акта за изкуствения интелект*, Brussels: European Commission, 2025.

въздействието върху защитата на данните. Той също така позиционира оценката на въздействието върху основните права като допълваща оценката на въздействието върху защитата на данните, като това може да включва и разглеждане на неравномерни въздействия между пациенти, диференцирани по биологичен и социален пол, както и по цвят на кожата и етнически произход. Освен това Европейското пространство за здравни данни има за цел да улесни сигурния достъп до здравни данни за предоставяне на здравни услуги и за научни изследвания, като същевременно поддържа високо равнище на защита на данните.¹⁶ Очаква се Европейското пространство за здравни данни да играе ключова роля за осигуряването на по-представителни набори от данни за разработването на биомедицински системи с ИИ, като подпомага оценяването, насочено към равнопоставеността, и намалява риска от недостатъчно представяне на пациенти с определени демографски характеристики.¹⁷

Накрая, по отношение на киберсигурността, Агенцията на ЕС за киберсигурност е изготвила насоки относно практиките за киберсигурност при ИИ. В тях се посочва, че съществуващите организационни стандарти (напр. системи за управление на сигурността и качеството по ISO) могат да бъдат адаптирани към контекста на ИИ.¹⁸ Насоките на Групата за координация по медицинските изделия относно киберсигурността за медицинските изделия очертават как производителите следва да изпълняват изискванията на MDR и IVDR, свързани с киберсигурността, чрез управление на риска и минимални изисквания за киберсигурност.¹⁹ Силното управление на киберсигурността е основа за равнопоставено внедряване на ИИ: защитата на целостта на данните, достъпа и надеждността на системите помага да се предотврати неволно засилване на неравенства, основани на биологичен и социален пол или цвят на кожата или етнически произход.

Взети заедно, тези правни инструменти създават сложна регулаторна среда, в която функционират системите с ИИ в здравеопазването. Предложеният тук регулаторен модел за ИИ в европейските обществени болници е изграден именно върху тези четири основни правни нива, които

¹⁶ European Union, [Регламент относно Европейското пространство за здравни данни \(EHDS\)](#), Brussels: European Union, 2024.

¹⁷ European Commission, [Европейско пространство за здравни данни \(EHDS\)](#), Brussels: European Commission, 2025.

¹⁸ European Union Agency for Cybersecurity (ENISA), [Киберсигурност на изкуствения интелект и стандартизация](#), Athens: ENISA, 2023.

¹⁹ Medical Device Coordination Group, [MDCG 2019-16 Rev. 1 – Насоки относно киберсигурността за медицински изделия](#), Brussels: European Commission, 2020.

уреждат управлението на здравния ИИ на практика. Съответно моделът предвижда механизми за наблюдение и смекчаване на несправедливите въздействия върху пациенти с различен биологичен и социален пол, цвят на кожата или етнически произход, през целия жизнен цикъл на ИИ.

Европейски и международни рамки за управление, приложими към ИИ в здравеопазването

Важно е да се отбележи, че разгледаните по-горе европейски регулаторни и стандартизационни рамки представляват основната правна основа за управлението на биомедицинския ИИ в рамките на Съюза. Наред с регулаторната среда на ЕС за биомедицински ИИ, съществуват и няколко международни необвързващи инструмента, които предоставят насоки за управление на рисковете, свързани с ИИ. Следва да се подчертае, че **допълнителните рамки за управление, разгледани в този раздел, са включени с цел да се илюстрират подходи за прилагане, които могат да бъдат прехвърлени в контекста на ЕС, като същевременно остават изцяло подчинени на европейските правни изисквания и принципите на основните права.** Те се използват селективно като аналогии от други правни системи, в които на практика са създадени механизми за управление през целия жизнен цикъл, по-специално в области като управление на риска на модели, алгоритмичен одит и институционална оценка на въздействието.

Например, Рамката за управление на риска при изкуствен интелект (AI Risk Management Framework – RMF) на Националния институт за стандарти и технологии (National Institute of Standards and Technology – NIST) предлага цялостен подход за управление на рисковете, свързани с ИИ, чрез подход, основан на жизнения цикъл и организиран около управление, картографиране на риска, измерване на риска и управление на риска.²⁰ Рамката изрично признава справедливостта и преодоляването на вредни системни пристрастия като основни измерения на надеждния ИИ. Освен това, в рамките на картографирането и измерването на риска, тя насърчава организациите да идентифицират и оценяват неравни въздействия върху пациенти, като се вземат предвид и различията по биологичен и социален пол, цвят на кожата и етнически произход. Целта е доверието и контролните механизми за управление на риска да бъдат интегрирани във всички етапи –

²⁰ National Institute of Standards and Technology, [Рамка за управление на риска при изкуствен интелект \(AI RMF 1.0\)](#), Gaithersburg, MD: NIST, 2023.



проектиране, внедряване и мониторинг, чрез съчетаване на техническа оценка с организационна отчетност. За болниците силата на Рамката за управление на риска при ИИ се състои в това, че тя може да бъде преведена в конкретни практики, като **ясно определени роли за надзор върху равнопоставеността, систематична оценка на представянето върху различни пациенти и документиращи стратегии за смекчаване на рисковете**, като същевременно се запазва по-широка перспектива, основана на правата.

От страна на системите за управление, стандартът на Международната организация по стандартизация ISO/IEC 42001 определя изисквания за организационна система за управление на ИИ, като се фокусира върху политики, цели, процеси и непрекъснато подобрене за отговорен ИИ.²¹ Макар да не предписва конкретни вреди, той подпомага интегрирането на принципите за справедливост и недискриминация в организационните политики и механизмите за контрол на риска, позволявайки на институциите да адресират потенциални системни пристрастия, свързани с биологичен и социален пол, цвят на кожата и етнически произход, по координиран начин. Допълващ го е ISO/IEC 23894, който предоставя насоки за интегриране на управлението на риска в разработването и внедряването на ИИ, включително идентифициране и адресиране на рискове, свързани с непреднамерени или непропорционални въздействия върху различни пациенти.²² За клиничния ИИ по-специално ISO 14971 определя принципи за управление на риска, широко използвани в регулацията на медицински изделия, които описват систематично идентифициране на опасности, оценка на риска, контрол на риска и мониторинг през целия жизнен цикъл.²³ Тези процеси могат да бъдат разширени така, че да включват неравна ефективност или резултати при пациенти с различни демографски характеристики като част от идентифицирането на вредите. Описаните стандарти са ценни, тъй като предлагат **структура на система за управление, която болниците могат да възприемат**. Те също така предоставят общ език, чрез който изисквания, свързани с равнопоставеност като оценка върху различни пациенти и смекчаване на системни пристрастия могат да бъдат ясно формулирани в обществените поръчки и договорите с доставчици.

²¹ International Organization for Standardization, *ISO/IEC 42001:2023 – Система за управление на изкуствения интелект*, Geneva: ISO, 2023.

²² International Organization for Standardization, *ISO/IEC 23894:2023 – Управление на риска при изкуствен интелект*, Geneva: ISO, 2023.

²³ International Organization for Standardization, *ISO 14971:2019 – Медицински изделия: прилагане на управление на риска при медицински изделия*, Geneva: ISO, 2019.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Етичните рамки за управление също предоставят важни нормативни насоки. Етичните насоки на Европейската комисия за надежден ИИ (EC's Ethics Guidelines for Trustworthy AI) представляват полезен, макар и необвързващ, инструмент на политиката.²⁴ Те определят седем основни изисквания за отговорни системи с ИИ, включително разнообразие, недискриминация и справедливост, които изрично изискват избягване на системни пристрастия и отчитане на въздействия върху хора с различни демографски характеристики, като биологичен и социален пол, цвят на кожата и етнически произход. Списъкът за оценка на надежден изкуствен интелект (Assessment List for Trustworthy Artificial Intelligence – ALTAI), доброволен инструмент за самооценка, разработен от Групата на високо равнище на експертите по ИИ към Европейската комисия, прилага тези насоки под формата на контролен списък за самооценка.²⁵ В рамките на регулаторен модел за здравеопазване стойността на Списъка за оценка на надежден изкуствен интелект се състои в неговия **формат, основан на въпроси, който може да бъде адаптиран в практически контролни списъци за обществени поръчки и внедряване**, насочващи заинтересованите страни да разглеждат въпроси като представителността на обучаващите данни, риска от несправедливи клинични резултати и адекватността на предпазните мерки срещу дискриминационни ефекти, като по този начин се интегрират съображения за равнопоставеност във взимането на решения от страна на болничния персонал.

От гледна точка на правата на човека, Рамковата конвенция на Съвета на Европа за изкуствения интелект представлява релевантен правен инструмент, изрично основан на принципите на човешко достойнство, прозрачност, отчетност, равенство и недискриминация и защита на личния живот, който предоставя правна основа за адресиране на дискриминационни ефекти, включително такива, свързани със системни пристрастия по пол, цвят на кожата и етнически произход в системите с ИИ.²⁶ По сходен начин Принципите на Организацията за икономическо сътрудничество и развитие относно ИИ²⁷ и Препоръката на ЮНЕСКО относно етиката на ИИ²⁸ предоставят международни нормативни насоки,

²⁴ European Commission, *Насоки за етични принципи за надежден изкуствен интелект*, Brussels: European Commission, 2019.

²⁵ European Commission, *Списък за оценка на надежден изкуствен интелект (ALTAI) – самооценка*, Brussels: European Commission, 2020.

²⁶ Council of Europe, *Рамкова конвенция относно изкуствения интелект (CAI)*, Strasbourg: Council of Europe, 2023.

²⁷ Organisation for Economic Co-operation and Development, *Принципи на ОИСР относно изкуствения интелект*, Paris: OECD, 2019.

²⁸ UNESCO, *Препоръка относно етиката на изкуствения интелект*, Paris: UNESCO, 2021.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI

които насърчават идентифицирането и смекчаването на несправедливи или неравномерни въздействия, като засилват очакването, че системите с ИИ не трябва да възпроизвеждат или задълбочават структурни неравенства. За етичното управление, специфично за здравеопазването, Световната здравна организация е формулирала шест основни принципа за ИИ в здравеопазването: автономност, благополучие, обяснимост, отговорност, справедливост и адаптивност.²⁹ **Принципът на справедливост изрично изисква внимание към различията в достъпа, ефективността и резултатите между различни пациенти, като се отчитат и фактори като биологичен и социален пол, цвят на кожата и етнически произход.**

Тези международни рамки за управление се съотнасят пряко с изискванията, произтичащи от Хартата, свързани с достойнството, равенството, недискриминацията и правото на здраве, и споделят общ акцент върху управлението през целия жизнен цикъл – принципи, които могат пряко да бъдат преведени в практически механизми за контрол в здравните институции.

Уроци от междусекторни регулаторни модели

Високорискови сектори като авиацията, банковото дело и фармацевтиката отдавна са разработили усъвършенствани системи за управление, насочени към контрол на технологичните опасности. Тези сектори предлагат ценни регулаторни подходи, които могат да бъдат адаптирани към управлението на ИИ в здравеопазването и към свързаните с него предизвикателства, свързани със системни пристрастия по пол, цвят на кожата и етнически произход.

Макар че системните пристрастия по пол, цвят на кожата и етнически произход не са технологични опасности в тесния смисъл на думата, те се проявяват като социотехнически рискове, когато данните, алгоритмите и организационните практики взаимодействат със структурни неравенства в здравеопазването. В този контекст „опасност“ се отнася до системното проявление на пристрастия в изходните резултати, решенията и клиничните изходи, а не до техния произход. Самото здравеопазване вече предоставя значителни доказателства за такива ефекти, включително документирані различия в диагностиката, лечението и предсказването на риска между пациенти с различни демографски характеристики, включително биологичен и социален пол, цвят на кожата и етнически произход. Следователно междусекторните примери допълват доказателствата от

²⁹ World Health Organization, *Етика и управление на изкуствения интелект в здравеопазването*, Geneva: WHO, 2021.

здравеопазването, като илюстрират механизми за управление на сложни системи, наред с емпирични данни от здравния сектор, които показват необходимостта от подобни регулаторни механизми.

Особено релевантен пример е Системата за управление на безопасността (Safety Management System – SMS), използвана в авиацията. Международните регулации в областта на гражданската авиация разглеждат безопасността като непрекъснат организационен процес, който позволява вземането на решения въз основа на идентифициране на опасности, управление на риска, гарантиране на безопасността (чрез мониторинг и одит), управление на промените и непрекъснато усъвършенстване.³⁰ Прилагането на тази логика към ИИ в здравеопазването би означавало **създаването на „Система за управление на безопасността на ИИ“, която изрично включва съображения за равнопоставеност**, в това число: (1) канал за докладване на опасности и почти настъпили инциденти от болничния персонал, чрез който да се регистрират потенциални случаи на системни пристрастия по пол, цвят на кожата или етнически произход; (2) официален регистър на риска за всяка ИИ система, документиращ различията в представянето на системата между различни пациенти; (3) механизми за контрол на промените, които оценяват въздействието на актуализации на модела, промени в данните (data drift) или промени в работните процеси върху резултатите, свързани с равенството; и (4) периодични прегледи за гарантиране на безопасността, придружени от коригиращи действия, насочени както към техническите рискове, така и към установените неравенства, свързани с пол, цвят на кожата или етнически произход. Изричният акцент на Акта за ИИ (чл. 9 и 72) върху управлението на риска през целия жизнен цикъл и наблюдението след пускане на пазара е в тясно съответствие с логиката на тази система за управление на безопасността.

Втори полезен подход идва от финансовото регулиране, където сложните количествени инструменти се разглеждат като „модели“, които изискват надеждна разработка, независима валидация и надзор върху управлението им. В банковия надзор сходни рамки за управление на риска са въведени в юрисдикции като Съединените американски щати (напр. Надзорните насоки на Федералния резерв на САЩ относно управлението на риска на модели – Supervisory Guidance on Model Risk Management) и предоставят пример за това как могат да бъдат организирани валидацията, надзорът и независимият

³⁰ International Civil Aviation Organization, [Ръководство за управление на безопасността](#), 4th ed., Montréal: ICAO, 2018.



преглед на моделите.³¹ Подобни подходи биха могли да бъдат приложени към биомедицинските системи с ИИ, за да се гарантира, че **инструментите за подпомагане на клиничните решения преминават независима валидация (вътрешна или външна), преди да бъдат внедрени в здравни заведения, като изрично се оценява представянето им при пациенти, разграничени по биологичен и социален пол, цвят на хората или етнически произход.** Минималният набор от документи за валидиране следва да съдържа информация за предназначението на системата, границите на нейната приложимост, известните режими на отказ и представянето ѝ при различни пациенти, включително установени различия и приложени мерки за тяхното смекчаване. Повторна валидация следва да се извършва винаги когато системата се внедрява в ново лечебно заведение, прилага се към различни групи пациенти, свързва се с ново оборудване или нови клинични протоколи, актуализира се от доставчика или показва признаци на влошаване на ефективността с течение на времето. Тези практики са в съответствие с Акта за ИИ (чл. 13, 14 и 60), който изисква тестване, прозрачност спрямо разпоредителите и човешки надзор, предназначен да открива и предотвратява системни пристрастия, произтичащи от прекомерно доверяване на автоматизирани решения, като по този начин гарантира, че подпомаганите от ИИ здравни грижи са едновременно безопасни и справедливи за различните групи от населението.

Що се отнася до управлението след внедряването, фармацевтичната регулация предоставя силен модел: във фармакологичната бдителност (pharmacovigilance) нежеланите реакции и сигналите за безопасност се събират и анализират систематично с цел откриване на нововъзникващи рискове. По-специално, Модул VI от Добрата практика по фармакологична бдителност на Европейската агенция по лекарствата (Good Pharmacovigilance Practices – GVP) определя подробни изисквания за събирането, управлението и подаването на съобщения за предполагаеми нежелани реакции, като отразява екосистемен подход към докладването на инциденти и последващите действия.³² **Сравнима система за „бдителност по отношение на ИИ“ (AI-vigilance) би могла да позволи на болниците да докладват и анализират инциденти, свързани с клинични решения, подпомагани от ИИ, с особено внимание към резултатите, свързани с равнопоставеността.** Основните елементи на такава система биха

³¹ Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, *Надзорни насоки за управление на моделния риск (SR 11-7)*, Washington, D.C., 4 April 2011.

³² European Medicines Agency, *GVP модул VI: Събиране и подаване на съобщения за подозирани нежелани реакции*, London: EMA, 2017

включвали: (1) систематично докладване, което обхваща всички наблюдавани неравенства в грижите или резултатите между различни групи от населението, разграничени по биологичен и социален пол или цвят на кожата и етнически произход; (2) задължително документиране на сериозни инциденти, при които системни пристрастия може да са допринесли за различно третиране; (3) задължения за уведомяване на доставчика и съответните органи при възникване на опасения, свързани с равнопоставеността (като органите се уведомяват, когато системни пристрастия водят до сериозна вреда, например погрешна диагноза или неравнопоставено лечение, докато рутинните неравенства се управляват в рамките на вътрешен контрол на качеството); и (4) функции за откриване на тенденции и сигнали, които идентифицират възникващи модели на системни пристрастия между отделения, лечебни заведения или в рамките на по-широката здравна мрежа. Такъв мониторинг би позволил проактивни мерки за смекчаване и непрекъснато подобрене, като гарантира, че подпомаганите от ИИ здравни грижи са не само безопасни, но и справедливи.

В допълнение, инструменти за управление в публичния сектор като оценки на въздействието на алгоритми (algorithmic impact assessments) и одити за системни пристрастия предоставят практически методи за оценка на обществения ефект на автоматизирани системи за вземане на решения, включително потенциални различия между групи от населението, разграничени по биологичен и социален пол или по цвят на кожата и етнически произход. Тези инструменти превеждат етичните принципи и стандартите за правата на човека в оперативни формати, които организациите могат да използват, за да предвиждат и смекчават неравенствата преди внедряването. Например Канадската федерална оценка на алгоритмичното въздействие (Algorithmic Impact Assessment) оценява рисковете на системите с ИИ и ги свързва с мерки за тяхното смекчаване.³³ Този инструмент изрично рамкира внедряването като проблем на управление, основано на въздействието. Освен това подходът на град Ню Йорк към автоматизираните инструменти за вземане на решения при заетост изисква одити за системни пристрастия и задължения за разкриване на информация към обществеността.³⁴ Съответно болници и граждански организации, работещи в здравеопазването, биха могли да въведат пакет „Оценка на въздействието и одит на ИИ“, който

³³ Treasury Board of Canada Secretariat, *Algorithmic Impact Assessment*, Ottawa: Government of Canada, 2026.

³⁴ New York City Department of Consumer and Worker Protection, *Автоматизирани инструменти за вземане на решения при заетостта*, New York: City of New York, 2023.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



комбинирана структурирана въпросна рамка, ориентирана към равнопоставеността, с периодични независими одити за справедливост, като предоставя публично достъпно резюме за представянето на системите в различните групи от населението и предприетите мерки за смекчаване. Прилагането на такъв пакет би допринесло за насърчаване на прозрачност, обществена отчетност и доверие на пациентите в здравните грижи, подпомагани от ИИ.

Включването на регулаторни модели от сектори като авиационната безопасност, управлението на моделни рискове във финансовия сектор, фармакологичната бдителност и алгоритмичното управление има за цел да представлява съзнателен трансфер на утвърдени принципи на управление. Тези области са подбрани, тъй като на практика прилагат формализирани подходи за откриване, наблюдение и смекчаване на рискове в сложни системи, при които могат да възникват различни въздействия върху отделни лица и групи от населението. Някои от тях вече включват практики, свързани със справедливостта, в оперативен смисъл, включително демографска стратификация на риска във фармакологичната бдителност, валидиране на модели и контрол на системни пристрастия във финансовия надзор, както и систематично докладване на нежелани събития, засягащи конкретни групи от населението. Тяхната релевантност се състои както в сходството в основната идея, така и в доказаното им приложение на управление през целия жизнен цикъл, задължително докладване и механизми за независим надзор, които могат директно да бъдат адаптирани за адресиране на рискове, свързани с равнопоставеността по отношение на биологичен и социален пол, цвят на кожата и етнически произход в биомедицинските системи с ИИ.

Схема за регулаторен модел на ИИ, съобразен с Хартата, за обществени болници и граждански организации

Този раздел представя архитектура на регулаторен модел, който е **основан на принципите на Хартата на ЕС, но е формулиран като практически процедури, които болниците и гражданските организации могат да прилагат** на практика. Хартата предоставя нормативната основа за управлението на здравните технологии в Европа, по-специално чрез принципите на човешко достойнство, равенство, недискриминация и

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



правото на здравеопазване.³⁵ Тези права са особено релевантни, когато автоматизирани системи влияят върху клинични решения, засягащи грижата за пациентите, включително потенциални различия между различни групи от населението, разграничени по биологичен и социален пол, цвят на кожата или етнически произход.

На най-фундаментално ниво чл. 1 от Хартата, който установява, че човешкото достойнство е ненакърнимо, подкрепя изискванията за управление, които гарантират, че инструментите с ИИ се използват по начин, който запазва човешката преценка и зачита автономността на пациента. На практика този принцип изисква **предпазни мерки срещу прекомерно разчитане на автоматизирани препоръки, механизми за оспорване на решения, подпомогнати от ИИ, както и структури за клиничен надзор, които предотвратяват системни пристрастия**, произтичащи от автоматизирано вземане на решения — тенденцията на хората, вземащи решения да се доверяват прекомерно на алгоритмични резултати, което може да задълбочи неравенствата, ако препоръките на ИИ системно поставят в неблагоприятно положение определени групи от населението.

По сходен начин чл. 20 и 21, които поставят фокус върху равенството пред закона и забраната за дискриминация, предоставят правната основа за изрични задължения за справедливост и смекчаване на системни пристрастия в биомедицинския ИИ. Болниците, които внедряват системи с ИИ, следва да **прилагат процеси за наблюдение на представянето на системите в различни групи от населението**, включително разграничени по биологичен и социален пол, цвят на кожата или етнически произход, както и в пресечни комбинации на тези характеристики. Когато бъдат установени неравенства в представянето или резултатите, следва да се прилагат коригиращи мерки чрез настройка на модела, промени в клиничните работни процеси или алтернативни пътища за вземане на решения, така че внедряването на ИИ да не задълбочава структурни неравенства.

Освен това чл. 35 от Хартата гарантира правото на високо равнище на защита на човешкото здраве и по този начин допълнително подкрепя принципа, че системи с ИИ могат да бъдат внедрявани само когато доказуемо допринасят за безопасно, ефективно и равнопоставено здравеопазване. ИИ технологиите не следва да създават нови бариери

³⁵ European Union, [Хартата на основните права на Европейския съюз](#), 2012/С 326/02, Strasbourg: EU.

пред достъпа до здравни услуги или да задълбочават съществуващи неравенства, засягащи исторически уязвими групи пациенти, като тяхното въздействие върху тези групи трябва да бъде непрекъснато оценявано.

За да се гарантират тези права на практика, предложеният в настоящия документ регулаторен модел възприема **подход на управление през целия жизнен цикъл, който структурира надзора върху ИИ около поредица от задължителни точки за вземане на решения или „контролни врати“**. Тези контролни врати отразяват регулаторната логика, заложена в Акта за ИИ и свързаното здравно законодателство, като същевременно предоставят на болниците ясна управленска пътека. Моделът приема, че болницата обикновено е разпоредител по смисъла на Акта за ИИ и доставчик на здравни услуги съгласно националното здравно право, но също така обхваща и хипотезата, при която болницата става доставчик/производител (например при вътрешно разработен ИИ).

Болниците са основни разпоредители и органи за прилагане и контрол на внедряването на системи с ИИ, докато гражданските организации действат като независими структури за надзор, одит и мониторинг, основан на правата, без оперативен контрол върху ежедневните клинични решения.

Прилагането на този регулаторен модел следва принципа на пропорционалност, като признава, че **болниците и гражданските организации могат да имат различни нива на технически, организационен и аналитичен капацитет в отделните държави членки и здравни системи**. Задължителни минимални изисквания се прилагат еднакво, за да се гарантира базово ниво на безопасност, прозрачност и предпазни мерки за равнопоставеност. Допълнително, по-напреднали практики като разширен анализ на резултатите по групи от населението, външен одит и по-сложни инструменти за мониторинг на системни пристрастия могат да бъдат въвеждани постепенно в зависимост от институционалния капацитет. Техническото приложение към документа също следва този подход, като прави разграничение между минимални и препоръчителни изисквания.

1. **Врата за обхват и класификация.** Преди възлагане на обществена поръчка или разработване институциите трябва да определят регулаторния статус на даден инструмент с ИИ. Това включва оценка дали системата представлява софтуер за медицинско изделие по смисъла на MDR или IVDR, дали представлява високорискова система с ИИ съгласно Акта за ИИ, или дали попада в по-нискорискова

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



категория, която все пак изисква управленски надзор. На този ранен етап болниците могат да маркират инструменти, чиято предназначена употреба може да има повишени последици за равнопоставеността или да засяга разнообразни групи от населението, като отбележат, че тези инструменти ще изискват по-задълбочена оценка на системни пристрастия по биологичен и социален пол и по цвят на кожата или етнически произход на по-късни етапи. Насоките, изготвени от Групата за координация по медицинските изделия относно квалификацията и класификацията на софтуера, предоставят практически примери, които могат да помогнат за избягване на погрешна регулаторна класификация и да положат основата за последващи оценки на системни пристрастия и равнопоставеност.³⁶ Болниците ръководят процеса по класификация и определяне на регулаторния статус, а гражданските организации предоставят независими консултативни становища относно рисковете за равнопоставеността, но нямат правомощия за вземане на решения на този етап.

2. **Врата преди възлагане на обществена поръчка и сключване на договор.** Болниците често разполагат с най-голям лост за влияние върху управлението на ИИ именно на етапа на обществените поръчки. На този етап институциите следва да гарантират, че доставчиците се ангажират договорно с адресиране на рисковете за равнопоставеност, включително системни пристрастия по пол, цвят на кожата и етнически произход, през целия жизнен цикъл на системите с ИИ. От доставчиците следва да се изисква предоставяне на документация, удостоверяваща съответствие с регулаторните изисквания, включително доказателства за извършени оценки на съответствието, резултати от валидиране на представянето и информация, необходима на разпоредителите за изпълнение на задълженията им по Акта за ИИ. Договорите за обществени поръчки следва изрично да разпределят отговорностите между доставчика и разпоредителя, включително изисквания, свързани с прозрачност, документация, човешки надзор и поддръжка след внедряването. Целта на този етап е да се интегрират осъзнатостта за системни пристрастия и отчетността в договорните задължения, като се създаде основа за безопасно и равнопоставено внедряване. Болниците носят отговорност за решенията по обществените поръчки и договорните отношения, а гражданските организации могат

³⁶ Medical Device Coordination Group, *Насоки относно квалификацията и класификацията на софтуер съгласно Регламент (ЕС) 2017/745 (MDR) и Регламент (ЕС) 2017/746 (IVDR)*, Brussels: MDCG, 2019.



да преглеждат документацията по поръчките от гледна точка на равнопоставеността и правата на човека и да предоставят необвързващи препоръки.

3. **Врата за валидиране.** Третият етап представлява врата за валидиране, която гарантира, че инструментите с ИИ се оценяват независимо преди да бъдат въведени в клиничните работни процеси. Валидирането следва да обхваща както общото представяне, така и резултатите при различни групи от населението, разграничени по биологичен и социален пол и по цвят на кожата и етнически произход. Този етап е критичен, тъй като биомедицинските системи с ИИ могат да показват неравномерна ефективност между различни групи от населението, когато обучаващите данни не са представителни или когато контекстуални различия между средата на разработване и средата на внедряване не са отчетени.³⁷ Оценката следва също така да документира всички приложени мерки за смекчаване и да подчертае оставащите рискове за равнопоставеността, като предоставя на болниците доказателствена основа за информирани решения за внедряване. Болниците носят отговорност за извършване или възлагане на валидирането и за интерпретацията на резултатите при вземането на решения за внедряване, а гражданските организации извършват независим преглед на представянето по групи от населението и рисковете от системни пристрастия, без да имат контрол върху резултатите от валидирането.
4. **Врата за внедряване.** След като системата успешно премине етапа на валидиране, следва четвъртият етап – вратата за внедряване. Болниците трябва да гарантират, че условията за работа подпомагат безопасното, равнопоставено и отговорно използване на системата. Това включва определяне на обучен персонал за осъществяване на надзор, внедряване на механизми, които позволяват на болничния персонал да отменя или деактивира алгоритмични препоръки, както и гарантиране, че потребителите разбират предназначението на системата, нейните ограничения и всички съображения, свързани с представянето ѝ при различни групи от населението. Болничният персонал следва да бъде информиран за потенциални неоправдани различия в резултатите, когато се отчитат характеристики като биологичен и социален пол или цвят на кожата и етнически произход, а пациентите следва да бъдат уведомявани, когато при

³⁷ Rajkomar, A., Hardt, M., Howell, M. D., Corrado, G., & Chin, M. H., [Осигуряване на справедливост в машинното обучение за напредък на здравната равнопоставеност](#), *Annals of Internal Medicine*, 169(12), 866-872, 2018.



предоставянето на медицинска помощ се използват инструменти за вземане на решения, подпомагани от ИИ. Болниците запазват пълната отговорност за решенията относно внедряването, интегрирането в клиничната практика и надзора върху прилагането, докато гражданските организации нямат роля в самото внедряване, но могат да преглеждат документацията за готовност за внедряване от гледна точка на равнопоставеността и правата на човека.

5. **Врата за мониторинг, одит и отчетност.** Тази последна врата регулира етапа след внедряването в жизнения цикъл на ИИ. Системите с ИИ следва да бъдат обект на непрекъснат мониторинг на цялостното им представяне, на различията между отделните групи от населението и на инцидентите, свързани с безопасността или равнопоставеността. Болниците следва да поддържат регистрационни файлове за работата на системите, да прилагат стандартизирани процедури за докладване на инциденти, които обхващат потенциални системни пристрастия по пол, цвят на кожата или етнически произход, и периодично да преоценяват представянето на системите чрез повторно валидиране. Когато възникнат неравенства или рискове от системни пристрастия, които не могат да бъдат ефективно смекчени, системата следва да бъде модифицирана, временно спряна или изведена от употреба. Констатациите следва да подпомагат както вътрешното управление, така и обучението и обмена на знания на ниво мрежа с цел предотвратяване на повтарящи се прояви на системни пристрастия. Болниците носят отговорност за текущия мониторинг, поддържането на регистрационни файлове, докладването на инциденти и прилагането на мерки за смекчаване, докато гражданските организации извършват независими одити, оценяват системните рискове за равнопоставеност в различни лечебни заведения и подпомагат ескалирането на устойчиво проявяващи се неравенства, без да упражняват оперативен контрол върху модификациите на системите.

Заедно тези врати на жизнения цикъл трансформират принципите на основните права в конкретни процеси на управление, способни да адресират рисковете от системни пристрастия по пол, цвят на кожата или етнически произход в клиничните системи с ИИ.

Фигура 1: Пътна карта за управление на ИИ, съобразена с Хартата (AEQUITAS)

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

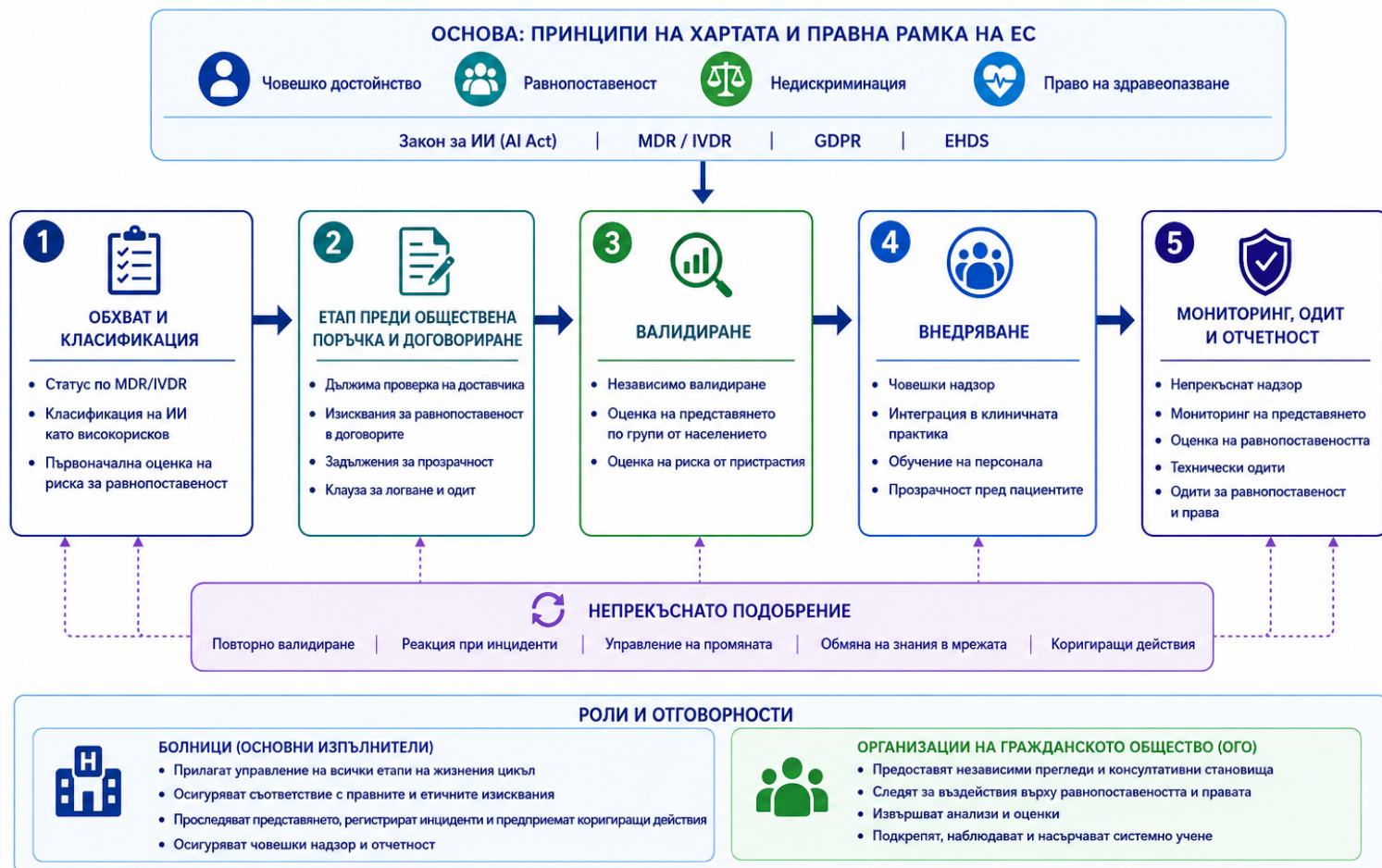
101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



ПЪТНА КАРТА ЗА УПРАВЛЕНИЕ НА ИИ, СЪОБРАЗЕНА С ХАРТАТА (AEQUITAS) ЗА БИОМЕДИЦИНСКИ ИИ В ПУБЛИЧНОТО ЗДРАВЕОПАЗВАНЕ



Практически политики, оперативни инструменти и управленски артефакти за професионалисти в болниците и гражданските организации

За да функционира ефективно в реални условия на здравеопазването, регулаторният модел, предложен в настоящия документ, предвижда разработването на практически инструменти, които могат да бъдат използвани повторно в различни институции и адаптирани към различни национални контексти. Моделът поставя акцент върху необходимостта от стандартизирани процедури, документация и образци на документи, които да позволят както на лечебните заведения, така и на гражданските организации последователно и съгласувано да внедряват и управляват системите с ИИ. Представеният по-долу набор от инструменти е съобразен

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



с изискванията на Акта за ИИ, насоките относно Регламента за медицинските изделия и Регламента за медицинските изделия за ин витро диагностика, както и с добрите практики за междуинституционално управление.

Етап на възлагане на обществената поръчка и сключване на договора

- **Пакет от документи за възлагане и договаряне.** Общността на ЕС, работеща в областта на обществените поръчки за системи с ИИ, разработи примерни договорни клаузи (за високорискови и за нискорискови системи), които впоследствие бяха актуализирани с оглед на изискванията на Акта за ИИ. Публичните органи са изрично насърчавани да използват тези клаузи, като към тях са предоставени и разяснения относно практическото им прилагане.³⁸ В контекста на използването на ИИ в здравеопазването, с цел оценка и ограничаване на риска от системни пристрастия, свързани с биологичния и социалния пол, цвета на кожата и етническия произход, тези клаузи могат да бъдат адаптирани като специфично приложение към договорите на лечебните заведения. Такова приложение следва да предвижда като минимум:
 - (1) Стандартизирано докладване на състава на обучаващите данни, включително представителството по биологичен и социален пол, цвят на кожата и етнически произход, с цел оценка на представителността на данните и риска от системни пристрастия;
 - (2) Разкриване на произхода на данните, при какви условия са достъпни и какви са последиците за прозрачността, независимото валидиране и оценката на системните пристрастия, свързани с биологичния и социалния пол, цвета на кожата и етническия произход;
 - (3) Предоставяне на информация за ефективността на системата при различните групи от населението, включително разграничени по биологичен и социален пол, цвят на кожата и етнически произход; документация за установени различия в представянето; описание на приложените мерки за ограничаване на системните пристрастия по време на разработването; както и доказателства от проведените валидиращи тестове, включително специфични за лечебното

³⁸ European Commission, [Updated EU AI Model Contractual Clauses Now Available](#), Brussels: European Commission, 2025.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



заведение съображения, свързани с равнопоставеността и справедливостта;

- (4) Изисквания за регистриране на действията и възможност за одит, които да гарантират, че системата може да генерира регистрационни записи, позволяващи провеждането на независими проверки за наличие на системни пристрастия или непропорционално неблагоприятно въздействие върху определени групи;
- (5) Задължения относно актуализациите и управлението на промените, изискващи доставчикът своевременно да уведомява лечебното заведение за актуализации на модела, повторно обучение или промени в работните процеси, както и да документира потенциалното им въздействие върху ефективността на системата по отношение на различните групи от населението;
- (6) Ангажменти за реакция при инциденти, свързани със системни пристрастия, задължаващи доставчика да уведомява лечебното заведение за установени грешки или неравнопоставено третиране, засягащи групи пациенти, разграничени по биологичен и социален пол или по цвят на кожата и етнически произход, както и да представи план за тяхното своевременно отстраняване;
- (7) Ясно разпределение на отговорностите, като се определят задълженията на доставчика и на внедряващата организация по отношение на спазването на изискванията на Акта за ИИ и управлението на рисковете, свързани с равнопоставеността и системните пристрастия.

Етапи на валидиране и преди внедряване

- **Политика за управление на здравните данни и осигуряване на представителност.** За ограничаване на системните пристрастия, свързани с биологичния и социалния пол, цвета на кожата и етническия произход, всички лечебни заведения и граждански организации, които внедряват системи с ИИ, следва да приемат официална политика, гарантираща, че данните, използвани за обучение, дообучаване или внедряване на системите, са представителни, пълни и точни за групите от населението, които обслужват. Това включва оценка на демографския състав на данните, документиране на недостатъчно представените групи пациенти и определяне на мерки за ограничаване на риска, когато това е необходимо. Политиката следва да обхваща както вътрешно

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



генерираните, така и външно предоставените набори от данни и да изисква проверка на резултатите, разпределени по групи от населението, както и на всички стратегии за ограничаване на системните пристрастия, приложени от доставчиците. Подобна рамка за управление е в съответствие с изискванията на Акта за ИИ (чл. 10), който изисква идентифициране, оценка и ограничаване на рисковете от системни пристрастия, и отразява утвърдените добри практики в научната литература относно ограничаването на системните пристрастия при използването на ИИ в здравеопазването през целия жизнен цикъл на системата – от събирането на данни до наблюдението след внедряването.³⁹ Тази политика сама по себе си не представлява отделен практически инструмент, а изпълнява ролята на общ управленски слой, който насочва разработването и прилагането на останалите инструменти. Тя служи като основополагащ принцип на етапа преди внедряването и осигурява връзката с процесите на непрекъснат мониторинг, одит и отчетност.

- **Базова рамка за киберсигурност и модел на заплахите, специфични за системите с ИИ.** Системите с ИИ, използвани в лечебните заведения, могат да бъдат уязвими не само към конвенционални кибератаки, като замърсяване на данните (data poisoning), създаване на подвеждащи входни данни (adversarial examples) или заобикаляне на модела (model evasion), но и към манипулации или грешки, които засилват системните пристрастия, свързани с биологичния и социалния пол, цвета на кожата и етническия произход. Това може да се прояви например чрез непропорционално неблагоприятно въздействие върху недостатъчно представени демографски групи в данните за обучение или при реалното използване на системата. Актът за ИИ (чл. 15) изисква системите да бъдат устойчиви на неразрешени промени в техните резултати или функциониране, което подчертава необходимостта от специфични за ИИ мерки за защита. Поради това лечебните заведения следва да разработят базова рамка за киберсигурност и модел на заплахите, специфични за системите с ИИ, който изрично отчита рисковете, свързани със системните пристрастия, наред с техническите заплахи. Този „жив“ документ изпълнява едновременно ролята на политика и практически инструмент. Той служи като контролен списък на етапа преди внедряването, който подпомага

³⁹ Char, D.S., Shah, N.H., & Magnus, D., *Прилагане на машинно обучение в здравеопазването — справяне с етичните предизвикателства*, *New England Journal of Medicine* 378: 981-983 (2018).



въвеждането на мерки за киберсигурност в съответствие с изискванията на Регламента за медицинските изделия, Регламента за медицинските изделия за ин витро диагностика, рамките за организационно управление на информационната сигурност (напр. ISO/IEC 27001), както и с подходите за моделиране на заплахите и планиране на тестове, отчитащи риска от системни пристрастия.⁴⁰ Документът подпомага също така непрекъснатия мониторинг след внедряването с цел своевременно откриване на нововъзникващи технически рискове и рискове, свързани със системните пристрастия, както и предприемането на навременни коригиращи действия. Чрез прилагането на тази политика лечебните заведения могат да повишат устойчивостта на своите системи, да гарантират равнопоставеност при вземането на решения, подпомагани от ИИ, и да осигурят съответствие с изискванията на Акта за ИИ (чл. 9–15), както и на MDR и IVDR.

- **„Досие на системата с ИИ“ за всяка отделна система.** Този инструмент представлява стандартизиран документ, който се изготвя и поддържа за всяка система с ИИ преди нейното внедряване в организация от системата на здравеопазването. Досието следва да обобщава предназначението на системата, начина на интегрирането ѝ в клиничните работни процеси, резултатите от валидирането и показателите за ефективност, като се обръща специално внимание на резултатите при пациентски групи, разграничени по биологичен и социален пол, цвят на кожата или етнически произход, както и на всички установени различия в представянето и приложените мерки за ограничаване на системните пристрастия. Документът следва също така да описва механизмите за човешки надзор, ясно да посочва кои решения остават изцяло в компетентността на медицинските специалисти и да съдържа информация за всички противопоказания, имащи значение за безопасността на пациентите и равнопоставеността при предоставянето на здравни услуги. Като предоставя формализиран и отчитащ риска от системни пристрастия преглед на системата с ИИ, досието подпомага прозрачността, вътрешното управление и регулаторния надзор, като гарантира, че решенията за внедряване изрично отчитат риска от неравнопоставени резултати за различните групи пациенти.

⁴⁰ Medical Device Coordination Group, [MDCG 2019-16 Rev. 1 – Насоки относно киберсигурността за медицински изделия](#), Brussels: European Commission, 2020.



- **Комбиниран пакет за оценка на въздействието върху основните права и оценка на въздействието върху защитата на данните.** Публичните здравни заведения, които внедряват високорискови системи с ИИ, все по-често са задължени да оценяват както рисковете за основните права, така и рисковете, свързани със защитата на личните данни. Обединяването на тези оценки в единен образец, който отговаря едновременно на изискванията на Акта за ИИ (чл. 27), Общия регламент относно защитата на данните и изискванията за оценка на въздействието върху защитата на данните, може значително да намали административната тежест, като същевременно гарантира систематичното разглеждане на ключови въпроси, свързани със засегнатите групи от населението, потенциалните вреди, мерките за ограничаване на риска и механизмите за подаване на жалби. Особено внимание следва да се отделя на потенциалните неравенства, които могат да засегнат пациентски групи, разграничени по биологичен и социален пол, цвят на кожата или етнически произход.

Комбинираният пакет за оценка на въздействието върху основните права и оценка на въздействието върху защитата на данните функционира основно като предварителен и периодично актуализиран одитен инструмент, разположен на етапа преди внедряването и в рамките на слоя за одит на равнопоставеността и основните права. Тези оценки представляват организирани и документиранни анализи на рисковете за основните права и защитата на личните данни и изпълняват ролята на контролни точки в управлението, а не на инструменти за непрекъснат мониторинг. Макар че следва да бъдат актуализирани при промяна на рисковете или условията, при които функционира системата, тяхната основна функция е да осигурят систематична и подлежаща на преглед оценка на потенциалните въздействия преди и по време на внедряването. По този начин те подпомагат спазването както на Акта за ИИ, така и на Общия регламент относно защитата на данните, както и идентифицирането и ограничаването на системните пристрастия, свързани с биологичния и социалния пол, цвета на кожата и етническия произход.

За да бъде укрепен аспектът, свързан със защитата на човешките права, при разработването на оценките на въздействието следва да се предвиди съдържателно участие на гражданските организации, така че да бъде използвана тяхната експертиза в областта на правата

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



на човека. Например Европейската мрежа на националните институции за правата на човека (ENNHRI) призовава Европейската комисия да допълни съответните образци с насоки и механизми за обществено консултиране, за да се гарантира ефективното прилагане на оценките на въздействието върху основните права, които отчитат равнопоставеността по отношение на биологичния и социалния пол, цвета на кожата и етническия произход.⁴¹

- **Политика за управление на промените.** Тази политика урежда актуализациите на системите с ИИ, включително повторно обучение на модели, софтуерни надграждания или изменения в клиничните работни процеси. Организацията следва ясно да дефинира какво представлява „съществена промяна“, която изисква повторна валидация или одобрение, като гарантират, че актуализациите не въвеждат непредвидени рискове за грижите за пациентите. По-специално, промените трябва да бъдат оценявани по отношение на потенциалното им въздействие върху равнопоставеността по пол, цвят на кожата и етнически произход, включително промени в резултатите, разпределени по групи от населението, или диференцирани резултати между групи от населението. Болниците следва също така да гарантират, че клиницистите и съответният персонал са информирани за изменения, които биха могли да повлияят на резултатите от грижите за различни групи пациенти, което да позволи подходящ надзор и мерки за смекчаване на риска. Чрез интегриране на съображенията за равнопоставеност в управлението на промените, болниците поддържат както безопасността, така и справедливостта на вземането на решения, подпомагано от ИИ в клиничната практика.
- **„Система за управление на качеството“ за внедрители, които същевременно са и доставчици.** Друг важен аспект на модела се отнася до ИИ, разработван вътрешно от болници или научноизследователски институции. Публичните здравни организации все по-често създават или адаптират собствени инструменти с ИИ, особено в академичните медицински центрове. В такива случаи институцията може на практика да действа като производител или доставчик, което въвежда допълнителни регулаторни отговорности. Европейските насоки относно

⁴¹ European Network of National Human Rights Institutions (ENNHRI), [Изявление относно гарантирането на ефективни оценки на въздействието върху основните права \(FRIA\) съгласно Акта на ЕС за изкуствения интелект](#), Brussels: ENNHRI, 2025.



медицинските изделия, разработени вътрешно, поясняват, че здравните институции следва да прилагат системи за управление на качеството и да поддържат техническа документация, демонстрираща, че вътрешно разработените инструменти отговарят на специфичните нужди на пациентите, които не могат да бъдат удовлетворени чрез търговски продукти.⁴² Системата за управление на качеството следва да гарантира, че резултатите от ИИ се оценяват по отношение на различията в ефективността между групи пациенти, разграничени по биологичен и социален пол, цвят на кожата или етнически произход, че стратегиите за смекчаване на рисковете са документирани и че механизмите за човешки надзор изрично запазват безопасността и справедливостта за пациентите. Дори когато се прилагат регулаторни изключения за вътрешно разработени медицински изделия, задълженията, произтичащи от основните права и по-широките рамки за управление на ИИ, остават напълно приложими, което прави Системата за управление на качеството ключов инструмент за интегриране на равнопоставеността в дизайна, валидирането и внедряването на биомедицински системи с ИИ.

Етап на мониторинг, одит и отчетност

- **„Доклад за ефективност и равнопоставеност по групи от населението“.** Тъй като вариациите в алгоритмичната ефективност могат да допринасят за неравенства в здравните резултати, болниците следва не само да изискват от доставчиците да предоставят съответната информация преди възлагане на обществена поръчка, но и да изготвят вътрешни анализи, показващи как системите се представят при пациенти с различни демографски характеристики. Тези анализи следва да се фокусират върху представяне по биологичен и социален пол, цвят на кожата или етнически произход, когато това е допустимо по закон и практически осъществимо, както и върху обосновка на всякакви неизбежни пропуски в представянето. Когато съществуват различия в представянето, организациите следва да документират стратегиите за тяхното смекчаване и да наблюдават ефективността на тези мерки във времето. Този подход отразява нововъзникващите очаквания в рамките на европейското регулиране на ИИ, според които

⁴² Medical Device Coordination Group, *MDCG 2023-1 – Насоки относно освобождаването за здравни институции (MDR/IVDR)*, Brussels: European Commission, 2023.



организациите активно оценяват и смекчават рисковете от системни пристрастия в наборите от данни и резултатите на моделите. Актът за ИИ (чл. 10) изрично изисква разглеждане и смекчаване на рисковете от системни пристрастия в наборите от данни и допуска ограничена обработка на специални категории данни за целите на откриване на системни пристрастия при строги гаранции.

Докладът за ефективност и равнопоставеност по групи от населението функционира основно като инструмент за непрекъснат мониторинг в рамката на управление. Той изисква текущо събиране на данни, редовно проследяване на представянето на системата при различни демографски групи и непрекъснато актуализиране на мерките за смекчаване на риска, когато бъдат установени диференцирани резултати. Това позволява на здравните организации да откриват и адресират рискове от системни пристрастия във времето. В същото време ключовите констатации от този доклад – като установени различия и ефективността на мерките за смекчаване – следва да бъдат обобщавани и включвани в одита за равнопоставеност и права (представен по-долу), като се гарантира прозрачност и обществена отчетност.

- **Двуслоен одитен режим, включващ технически одит и одит за равнопоставеност и права.**

Осъществим одитен режим за болниците следва да балансира между строгост и практическа приложимост и да бъде разбираем както за персонала, така и за обществеността. В този контекст одитът представлява формализиран, периодичен преглед на съответствието, ефективността и въздействието на една система с ИИ върху правата, предназначен да допълва текущия мониторинг, а не да го заменя.

Технически одит: този слой представлява поверителен оперативен преглед и обхваща детайлни проверки, които остават вътрешни, включително тестове за сигурност, оценки на устойчивостта, клинична валидация на представянето и съответствие с изискванията за управление на данните. Особено внимание следва да се отделя на представянето при различия по биологичен и социален пол, цвят на кожата и етнически произход, като се оценява дали изходите на модела системно поставят в неравностойно положение определени пациенти. Тези дейности следва да бъдат в съответствие със

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



съответните разпоредби на Акта за ИИ (чл. 9–15), както и MDR и IVDR, когато е приложимо.

Одит за равнопоставеност и права: този втори слой се фокусира върху прозрачност, отчетност и смекчаване на системни пристрастия. Той следва да документира неравенствата в резултатите, популациите, включени в тестването, известните ограничения и механизмите за подаване на жалби, с изричен анализ на въздействията по отношение на биологичен и социален пол, цвят на кожата и етнически произход. Следва също така да бъде отразено доказателство, че човешкият надзор е реално ефективен и способен да коригира пристрастни резултати. В съответствие с подхода за одит и разкриване на системни пристрастия, използван в регулацията на заетостта,⁴³ обобщение, предназначено за широката общественост, може да комуникира ключовите констатации, мерките за смекчаване и въведените подобрения, без да разкрива собствена или поверителна информация. Комбинираният пакет от оценка на въздействието върху основните права и оценка на представителността на данните също се ситиуира в слоя на одита за равнопоставеност и права, като засилва строгата оценка на основните права, защитата на данните и демографската равнопоставеност.

- **„Регистрационен запис за бдителност при ИИ“.** Мониторингът след внедряване също се подпомага чрез създаването на лог за бдителност при ИИ, вдъхновен от системите за фармаконадзор, използвани при регулирането на лекарствата.⁴⁴ Здравните специалисти следва да имат възможност да докладват инциденти, „почти инциденти“ или неочаквани резултати, свързани с решения, подпомогнати от ИИ, като се включват изрични полета за регистриране на потенциални наблюдавани различия по биологичен и социален пол, цвят на кожата или етнически различия. Докладите следва да се анализират систематично за възникващи модели, включително диференцирано въздействие върху конкретни демографски групи, което да позволява на организациите да откриват сигнали за безопасност, проблеми с равнопоставеността или

⁴³ New York City Department of Consumer and Worker Protection, [Автоматизирани инструменти за вземане на решения при заетостта \(AEDT\)](#), New York: City of New York, 2025.

⁴⁴ European Medicines Agency, [GVP модул VI – Събиране и подаване на съобщения за подозирани нежелани реакции \(Rev 2\)](#), London: EMA, 2017.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



тенденции на системни пристрастия. Когато е необходимо, а особено при инциденти, показващи неравнопоставено третиране или риск за защитени групи, те следва да бъдат съобщавани своевременно на доставчиците на системата, съответните болнични управленски комитети и регулаторните органи. Регистрационният запис следва също така да подхранва периодичните одити за равнопоставеност и права, като по този начин подпомага непрекъснатото подобрене и смекчаването на системни пристрастия при внедряването на клиничен ИИ.

- **Сътрудничество на мрежово ниво и споделяне на инциденти.** Европейската мрежа от граждански организации и обществени болници, разработена в рамките на AEQUITAS, функционира като споделен слой за управление, предоставящ обща терминология и стандарти за докладване на инциденти и констатации, свързани със системни пристрастия по отношение на пол и раса в системите с ИИ. Тази мрежа подпомага междуинституционалното обучение и разпространението на добри практики за откриване, смекчаване и коригиране на неравнопоставености, по аналогия с подходи в секторите с високи изисквания за безопасност и фармаконадзора. Чрез споделяне на данни за инциденти, анализ по групи от населението и стратегии за смекчаване на рискове, мрежата позволява на болниците и гражданските организации да идентифицират системни модели на неравнопоставеност и да координират целенасочени интервенции. Тъй като Европейското пространство за здравни данни улеснява както първичното използване (грижи за пациентите), така и вторичното използване (изследвания и политики) в рамките на регулирана система, мрежата може да съгласува своите практики за споделяне на данни и обучение с графиците за прилагане и изпълнителните актове по Европейското пространство за здравни данни, докато те се развиват, като по този начин подпомага по-равнопоставено и прозрачно внедряване на системи с ИИ между институциите.
- **Обучителен модул за болничен персонал и експерти от граждански организации.** С цел смекчаване на автоматизационното пристрастие и неравнопоставеностите при клинични системи с ИИ, регулаторният модел AEQUITAS включва задължително обучение, предназначено да подготви здравните специалисти с умения да разбират ограниченията на системите с ИИ, да интерпретират резултатите критично и да отменят или спират системите, когато

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 – AEQUITAS – CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз



изходите могат да доведат до пристрастни или неравноправни резултати. Обучението акцентира върху осведомеността относно рисковете от системни пристрастия по пол и раса, принципите на основните права и етичните съображения, включително равнопоставеност, недискриминация и включване на пациентите.

Този модул може да се разглежда като подпомагащ инструмент, който активно поддържа прилагането на регулаторния модел, като гарантира, че персоналът, отговорен за надзора върху ИИ, може компетентно да изпълнява процедури по мониторинг, валидация и отчетност през призмата на равнопоставеността. Обучението включва казусно базирани упражнения за откриване и смекчаване на различия по биологичен и социален пол, цвят на кожата и етнически различия.

Заклучение

ИИ предлага значителни възможности за подобряване на здравните резултати, но без внимателно управление той може да възпроизвежда или задълбочава неравнопоставености по биологичен и социален пол, цвят на кожата или етнически произход, както и други форми на системни пристрастия. Регулаторният модел, предложен в настоящия документ, превежда принципите на основните права в конкретни гаранции през целия жизнен цикъл на ИИ — от обществените поръчки и валидацията до внедряването, мониторинга и отчетността, с изрично внимание към откриването и смекчаването на различия в резултатите между различни групи пациенти. Чрез прилагане на архитектура за управление на ИИ, структурирана около ясно определени етапи на жизнения цикъл, политики, практически инструменти и артефакти, болниците и гражданските организации могат систематично да оценяват, смекчават и документират рискове, включително неравнопоставени въздействия върху маргинализирани групи пациенти. Чрез интегриране на междусекторен опит, изискванията на правото на ЕС и международни рамки за управление, моделът подпомага здравните специалисти в отговорното прилагане на системи с ИИ и гарантира равнопоставена, безопасна и прозрачна здравна грижа за всички пациенти, независимо от биологичен и социален пол, цвят на кожата, етнически произход или други характеристики.

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA. Номер на проекта:

101215009 — AEQUITAS — CERV-2024-CHAR-LITI



Съфинансирано от
Европейския съюз

