



# Генеральная Ассамблея

Distr.: General  
17 July 2024  
Russian  
Original: English

---

## Семьдесят девятая сессия

Пункт 71 b) предварительной повестки дня\*

**Поощрение и защита прав человека: вопросы прав человека, включая альтернативные подходы в деле содействия эффективному осуществлению прав человека и основных свобод**

## Права человека и международная солидарность

### Записка Генерального секретаря

Генеральный секретарь имеет честь препроводить Генеральной Ассамблее доклад Независимого эксперта по вопросу о правах человека и международной солидарности Сесилии М. Байе, представленный в соответствии с резолюцией [53/5](#) Совета по правам человека.

---

\* [A/79/150](#).



**Доклад Независимого эксперта по вопросу о правах человека и международной солидарности  
Сесилии М. Байе**

**Искусственный интеллект и международная солидарность —  
на пути к разработке ориентированных на человека систем  
искусственного интеллекта с учетом принципа  
международной солидарности**

*Резюме*

Настоящий доклад является первым докладом, подготовленным для Генеральной Ассамблеи Независимым экспертом по вопросу о правах человека и международной солидарности Сесилией М. Байе. Он представляется в соответствии с резолюцией 53/5 Совета по правам человека. Независимый эксперт рассматривает существующую в настоящее время проблему усиления глобального неравенства и объясняет, как с помощью стратегии разработки систем искусственного интеллекта с учетом принципа международной солидарности можно прояснить обязанности государств, корпораций и гражданского общества по расширению равного доступа к технологиям и обеспечению участия групп и лиц, находящихся в уязвимом положении, на недискриминационной основе. В настоящем докладе Независимый эксперт предлагает добиваться того, чтобы управление искусственным интеллектом подкрепляло государственные и корпоративные процедурные механизмы должной осмотрительности для привлечения прямых и косвенных заинтересованных сторон к обработке данных и принятию решений на протяжении жизненного цикла систем искусственного интеллекта.

## I. Введение

1. Международная солидарность — это принцип международного права и универсальная ценность, которой уделяют внимание политики, занимающиеся решением глобальных задач, связанных, в частности, с пандемиями, изменением климата, неравенством и чистой энергией<sup>1</sup>. В мире зарождается новое сознание, ориентированное на трансформационный переход от неолиберального неравенства к бережному отношению, солидарности и коллективным действиям посредством революции ценностей<sup>2</sup>. Необходимо, чтобы искусственный интеллект (ИИ) использовался для того, чтобы объединять человечество, а не разделять его. В соответствии со стандартами в области прав человека системы ИИ должны содействовать проведению коллективных международных акций солидарности в разных странах<sup>3</sup>. В 2022 году Организация Объединенных Наций по вопросам образования, науки и культуры (ЮНЕСКО) выпустила Рекомендацию об этических аспектах искусственного интеллекта, в которой признается необходимость укрепления солидарности для облегчения справедливого доступа к технологиям на основе ИИ и решения возникающих в связи с ними проблем в отношении разнообразия и взаимосвязанности культур и этических систем. Кроме того, концепция солидарности может способствовать смягчению последствий потенциального злоупотребления ИИ и полной реализации потенциала ИИ и обеспечивать, чтобы в основе национальных стратегий в области ИИ лежали этические принципы.

2. Как отмечается в одном из исследований, системы ИИ все чаще внедряются в странах, где создаются «умные города», чтобы облегчить отслеживание и мониторинг граждан — возможно, во всех аспектах их жизни — в условиях, когда все чаще используется биометрическая регистрация, а также в качестве инструмента международного развития и гуманитарной помощи при перемещении населения и в других кризисах<sup>4</sup>. Ожидается, что государственные учреждения и сторонние организации (включая компании и организации гражданского общества) по всему миру будут гораздо активнее использовать ИИ, чтобы рассматривать дела и иски и принимать решения по ним, используя персональные данные. Наблюдение с помощью ИИ особенно затрагивает малоимущие слои населения, поскольку государственные учреждения используют ИИ для выявления случаев нецелевого использования социальных пособий, в качестве инструмента для осуществления «чрезмерного полицейского контроля» за маргинализированными районами, а также в контексте отслеживания нелегальной миграции<sup>5</sup>. Невозможно переоценить, насколько важно признавать межсекторальные факторы уязвимости для дискриминации с помощью ИИ, в том числе такие как расовая, этническая и гендерная принадлежность, религиозные убеждения, место проживания, национальность и социально-экономический статус<sup>6</sup>.

3. Более того, в докладе «Индекс данных, касающихся искусственного интеллекта» 2024 года сделан вывод, что «разработчики ИИ не обеспечивают достаточный уровень транспарентности, особенно в части раскрытия используемых для обучения данных и методик. Из-за этого недостатка транспарентности

<sup>1</sup> См. Global Nation, *Global Solidarity Report 2023* (September 2023).

<sup>2</sup> Kurt April, “AI-Induced Solidarity Economy: The Need for Stewardship Orientation” in *Effective Executive*, vol. 26, No. 3 (2023).

<sup>3</sup> Материалы, представленные Форумом по вопросам управления Интернетом. См. также <https://intgovforum.org/en/content/pnai-report>.

<sup>4</sup> Linnet Taylor, “What is data justice? The case for connecting digital rights and freedoms globally”, *Big Data and Society*, vol. 4, No. 2 (July–December 2017).

<sup>5</sup> Ibid.

<sup>6</sup> Ibid.

затрудняются усилия по дальнейшему изучению надежности и безопасности систем ИИ»<sup>7</sup>. Кроме того, компании, работающие в сфере информационно-коммуникационных технологий (ИКТ), в большей степени сотрудничают с организациями гражданского общества из стран глобального Севера, что может привести к ослаблению солидарности с организациями гражданского общества из стран глобального Юга<sup>8</sup>. Устранение этого дисбаланса осложняется тем, что группы из стран глобального Юга имеют ограниченный доступ к сетям, которые занимаются вопросами подотчетности ИИ за нарушения прав человека.

4. При подготовке доклада Независимый эксперт обратилась с просьбой предоставить материалы к государствам-членам, организациям гражданского общества, компаниям и ученым. В июле 2024 года была проведена консультация с представителями гражданского общества. Был проведен обзор научной литературы, докладов договорных органов Организации Объединенных Наций и уставных органов, а также анализ международных, региональных и национальных правовых стандартов.

## **II. Критерии разработки ответственного искусственного интеллекта с учетом принципа солидарности: прозрачность, справедливость, недискриминация и инклюзивность**

5. Учитывая, что полномочия по обработке данных и принятию решений все чаще делегируются ИИ, существует фундаментальная необходимость в обеспечении прозрачности, чтобы гарантировать соблюдение принципов справедливости, недискриминации и инклюзивности. В докладе «Индекс данных, касающихся искусственного интеллекта» 2024 года справедливость определяется как «создание справедливых алгоритмов, недопущение предвзятости или дискриминации и учет различных потребностей и обстоятельств всех заинтересованных сторон, что позволяет соответствовать более широким общественным стандартам справедливости»<sup>9</sup>. В этом докладе также отмечается следующее: «Согласно результатам, хотя большинство компаний полностью внедрили хотя бы одну меру справедливости, по-прежнему не обеспечена комплексная интеграция. В среднем по миру было принято 1,97 меры по обеспечению справедливости из пяти», причем речь идет о тех мерах, данные по которым были собраны в рамках исследования «Глобальное положение дел в отношении ответственного искусственного интеллекта». Всемирный союз сравнительного анализа отметил, что по состоянию на 2023 год лишь четверть из 200 самых влиятельных технологических компаний мира соблюдают минимальный стандарт раскрытия информации о том, какие этические принципы, имеющие отношение к цифровой инклюзивности, они соблюдают в области ИИ<sup>10</sup>.

6. В «Римском призыве к обеспечению этических аспектов ИИ» принцип инклюзивности отмечен как один из ключевых, а на мероприятии «Этические аспекты ИИ во имя мира», состоявшемся в Хиросиме в июле 2024 года, председатель Форума мира в Абу-Даби шейх Абдалла бен Байя подчеркнул, что «чтобы системы

<sup>7</sup> Institute for Human-Centered AI, Stanford University, *Artificial Intelligence Index Report 2024* (Stanford University, California, 2024).

<sup>8</sup> См. <https://www.business-humanrights.org/en/from-us/briefings/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/>.

<sup>9</sup> *Artificial Intelligence Index Report 2024*.

<sup>10</sup> См. <https://www.worldbenchmarkingalliance.org/impact/investor-statement-for-ethical-ai-2024/>.

и продукты на основе ИИ были не только технически продвинутыми, но и правильными с точки зрения морали, необходимо при решении вопросов, касающихся развития искусственного интеллекта, которое сопряжено с разными интересами, вредными и полезными факторами, соблюдать принципы сотрудничества, солидарности и совместной работы»<sup>11</sup>. В проекте Глобального цифрового договора от 26 июня 2024 года сформулирована специальная инклюзивная цель: «Наше сотрудничество позволит преодолеть цифровой разрыв внутри государств и между ними и создать цифровую среду, способствующую поощрению и обеспечению разнообразия»<sup>12</sup>.

7. В соответствии с Руководящими принципами предпринимательской деятельности в аспекте прав человека технологическим компаниям следует проводить оценку воздействия на права человека и выявлять риски в рамках процессов контроля качества, а также взаимодействовать с внешними заинтересованными сторонами при оценке положения в области прав человека<sup>13</sup>. Можно предположить, что в рамках подхода к развитию ИИ с учетом принципа международной солидарности особое внимание будет уделяться необходимости создания конкретных процедурных подходов для борьбы с дискриминацией и укрепления инклюзивности. Рекомендуется интегрировать принцип солидарности между поколениями в нормы регулирования ИИ<sup>14</sup>.

8. Можно использовать концепцию социальной солидарности для распределения выгод и издержек от внедрения ИИ, для поощрения разнообразия в рамках траекторий развития ИИ и для обеспечения транспарентности и соблюдения норм, чтобы устранить информационную асимметрию, касающуюся ИИ<sup>15</sup>. Независимый эксперт подчеркивает, что важно учитывать аспекты солидарности при оценке процессов разработки и внедрения ИИ.

#### **А. Инклюзивность и алгоритмическая недискриминация как цели, сопряженные с развитием искусственного интеллекта на основе принципа солидарности**

9. Необходимость регулирования ИИ для устранения предвзятости и дискриминации, а также для обеспечения безопасности имеет огромное значение. ИИ может способствовать эмансипации путем предоставления образовательных возможностей в удаленных районах, обеспечивает доступность материалов на различных языках благодаря функции перевода, а также используется для борьбы со стереотипами и языком ненависти, поэтому необходимо постоянно поддерживать исследования, посвященные социальным последствиям использования ИИ. Учреждениям, разрабатывающим ИИ в рамках процессов принятия решений (в том числе в судебной и административной сферах), следует быть инклюзивными и стремиться привлекать уязвимые группы и ориентированные на обеспечение демократии гражданские субъекты на всех этапах — от

<sup>11</sup> См. <https://www.romecall.org/> и <https://www.romecall.org/ai-ethics-for-peace-hiroshima-july-9th-2024/>.

<sup>12</sup> См. проект договора от 26 июня на сайте <https://www.un.org/techenvoy/global-digital-compact>.

<sup>13</sup> Kate Jones, *AI Governance and Human Rights: Resetting the relationship*, Research Paper, International Law Programme (London, Royal Institute of International Affairs, 10 January 2023).

<sup>14</sup> Sébastien Fassiaux, “Preserving consumer autonomy through European Union regulation of artificial intelligence: a long-term approach”, *European Journal of Risk Regulation*, vol. 14, special issue No. 4 (December 2023). См. также Jon Truby and others, “A sandbox approach to regulating high-risk artificial intelligence applications”, *European Journal of Risk Regulation*, vol. 13, No. 2 (June 2022), recognizing solidarity as a principle within European Union AI regulation.

<sup>15</sup> Juan C. Mateos-García, “The Complex Economics of Artificial Intelligence” (2 December, 2018) (unpublished working paper).

планирования до использования — с целью предотвращения нарушений прав человека и смягчения вреда. Кроме того, важно создать независимый механизм надзора для защиты данных, способный регулировать сбор и обработку персональных данных на различных этапах планирования и внедрения ИИ и разрабатывать соответствующие рекомендации. Независимый эксперт поддерживает модель условного демографического неравенства, которая служит стандартным набором статистических данных для случаев дискриминации, связанной с автоматизированными системами<sup>16</sup>.

10. В одном из материалов, представленных для настоящего доклада, излагается концепция инклюзивного ИИ, предусматривающая реализацию принципа международной солидарности путем создания платформ, которые распространяют точки зрения традиционно недопредставленных групп населения в глобальном диалоге, например переводческие платформы на основе ИИ и платформы социальных сетей, используемые активистами для налаживания контактов, а также путем расширения возможностей гражданского общества для привлечения правительств к ответственности, например, в случаях коррупции<sup>17</sup>.

## **В. Национальные стандарты инклюзивности и недискриминации**

11. На национальном уровне есть множество стандартов (большинство из них носят рекомендательный, необязательный характер или существуют на стадии проекта), в которых отмечается приверженность принципам равенства, инклюзивности и недискриминации, однако отдельная институциональная компетенция и конкретные процедурные механизмы, обеспечивающие соблюдение таких стандартов и предоставление средств правовой защиты в случае нарушений, часто отсутствуют.

12. В Австралии существуют рекомендательные, не имеющие обязательной силы этические принципы ИИ, в которых признается, что системы ИИ должны приносить пользу людям, обществу и окружающей среде, а также быть основаны на идее уважения прав человека, разнообразия и автономии отдельных людей<sup>18</sup>. В Китае также приняты временные меры по регулированию предоставления услуг на основе генеративного искусственного интеллекта, предусматривающие предотвращение дискриминации по признаку этнического, географического и гендерного происхождения, религии, национальности, возраста, рода занятий и состояния здоровья, а также осуществление деятельности по обеспечению большего разнообразия и недискриминационного характера используемых для обучения данных<sup>19</sup>. В Индии разработана национальная стратегия в области искусственного интеллекта, в которой рассматриваются вопросы равенства и недискриминации<sup>20</sup>. В этой стратегии отмечается, что системы ИИ должны беспристрастно относиться к лицам, которые находятся в одинаковых обстоятельствах, имеющих отношение к принятию решения, что системы ИИ не должны

<sup>16</sup> Sandra Wachter, Brent Mittelstadt and Chris Russell, “Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI”, *Computer Law and Security Review*, vol. 41 (2021).

<sup>17</sup> Материалы, представленные Джейком Окечку Эффоду, доцентом Школы права имени Линкольна Александра, Метрополитенский университет Торонто, Канада.

<sup>18</sup> См. <https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>.

<sup>19</sup> См. <https://www.chinalawtranslate.com/en/generative-ai-interim/>.

<sup>20</sup> См. <https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf>; <https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf>; и <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>.

отказывать квалифицированному лицу в возможности по причине его принадлежности к той или иной группе и что они не должны увеличивать вредные исторически сложившиеся и социальные различия на основе религии, расы, касты, пола, происхождения, места рождения или проживания в таких вопросах, как образование, трудоустройство или доступ в общественные места. Кроме того, в этой стратегии определена задача не допускать несправедливого лишения возможности пользоваться услугами и льготами. Бразилия предложила регламент об ИИ, в котором предусмотрен стандарт в отношении недискриминации, включающий право оспаривать решения и требовать вмешательства со стороны человека, а также устранения прямых, косвенных, незаконных или оскорбительных дискриминационных предрассудков<sup>21</sup>. В Испании создан независимый орган по вопросам равного обращения и недискриминации, который контролирует и поощряет использование этических, надежных и соблюдающих основные права человека систем ИИ<sup>22</sup>.

13. В Соединенных Штатах Америки существует документ «Законопроект о правах ИИ: заставим автоматизированные системы работать на благо американского народа», который подкрепляет указ, содержащий призыв к независимой оценке и ведению отчетности для обеспечения недискриминации<sup>23</sup>. В Кении существует проект кодекса практических мер в отношении ИИ, в котором фактически рекомендуется создать механизм для внедрения стандартов недискриминации, в том числе путем фиксирования оценок справедливости, ведения учета шагов, предпринятых для устранения предвзятости, и документирования стратегий, направленных на предотвращение дискриминационных результатов<sup>24</sup>.

### **С. Риски, связанные с использованием систем искусственного интеллекта с элементами предвзятости**

14. Было признано, что «недорогой и осознанный доступ к Интернету стал фундаментальной необходимостью для полной реализации всех прав человека и основных свобод и обеспечения демократии, развития и социальной справедливости»<sup>25</sup>. Это привело к появлению целого ряда частных инициатив, которые посвящены вопросам цифровых прав и направлены на решение проблем инфраструктуры Интернета, его работы и его использования с точки зрения повышения доступности. Из-за предвзятости ИИ отдельные люди или группы людей подвергаются стереотипному представлению или предрассудкам, когда их сравнивают с другими людьми с помощью алгоритмов ИИ. Предвзятость ИИ проявляется при разработке системы ИИ, сборе и интерпретации данных, а также при прямом и косвенном взаимодействии с заинтересованными сторонами. Специальный докладчик по вопросу о современных формах расизма, расовой дискриминации, ксенофобии и связанной с ними нетерпимости подтвердила, что «проблемы с данными, проблемы разработки алгоритмов, намеренное дискриминационное использование искусственного интеллекта и вопросы подотчетности» по-прежнему вызывают трудности (см. [A/HRC/56/68](#)). В докладе ЮНЕСКО об

<sup>21</sup> См. законопроект № 2338 от 2023 года. Доступен по ссылке <https://www.senado.leg.br/web/atividade/materias/-/materia/157233>.

<sup>22</sup> Материалы, представленные Испанией.

<sup>23</sup> См. <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>; см. также <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>.

<sup>24</sup> См. [https://www.dataguidance.com/sites/default/files/kebs-tc\\_094\\_n66\\_public\\_review\\_kenya\\_standard\\_dks\\_3007\\_ai\\_code\\_of\\_practice.pdf](https://www.dataguidance.com/sites/default/files/kebs-tc_094_n66_public_review_kenya_standard_dks_3007_ai_code_of_practice.pdf).

<sup>25</sup> George A. Walker, “Technology Law, Rights and Ethics — One Choice, One Future”, *The International Lawyer*, vol. 56, No. 1 (2023).



ИИ за 2019 год отмечается, что солидарность является одним из элементов создания «общества, основанного на знаниях», и в нем также объясняется настоятельная необходимость борьбы с дискриминацией в области ИИ, возникающей в результате намеренного или ненамеренного дискриминационного программирования, наличия элементов предвзятости в данных, используемых для алгоритмов машинного обучения, или других факторов<sup>26</sup>:

Многие виды дискриминации могут иметь не прямой характер; к примеру, алгоритм, основанный на анализе использования мобильной связи для определения кредитоспособности человека, является дискриминационным, если он присваивает высокий кредитный риск женщинам в сообществах i) редко пользующихся мобильной связью и ii) не имеющих мобильных телефонов. Применяемое условие может показаться равным и справедливым, но оно ставит в невыгодное положение определенную группу. Алгоритмы могут как обуславливать, так и усугублять эти различные формы дискриминации. Существующие социальные и политические предрассудки систематизируются в алгоритмах машинного обучения множеством способов. Более того, следует подробнее изучить потенциально новые формы дискриминации, которые могут возникнуть из-за использования технологий ИИ, например, когда решения об исключении принимаются на основе статистических корреляций, которые не обязательно соответствуют социально важным характеристикам, но тем не менее имеют тесную связь с личностью человека.

15. В докладе «Индекс данных, касающихся искусственного интеллекта» 2024 года описывается, как токенизация языка (при которой язык разбивается на компоненты для анализа) отрицательно сказывается на незападных языках, поскольку с такими языками ИИ будет работать менее эффективно, чем с западными языками<sup>27</sup>. Существуют также организации гражданского общества, которые непосредственно занимаются проблемой алгоритмической предвзятости, исключения и дискриминации в отношении расовых, этнических или религиозных меньшинств, а также по признаку гендерной принадлежности или других факторов идентичности и работают над повышением доступности<sup>28</sup>. Организация «Международная амнистия» и Национальная ассоциация содействия прогрессу «цветного» населения выразили обеспокоенность по поводу инструментов прогностической деятельности полиции<sup>29</sup>, использования автоматизированных систем для определения доступа к медицинским и социальным услугам, наблюдения за перемещениями беженцев и мигрантов, а также необъективного воздействия на расовые меньшинства инструментов ИИ, используемых для распознавания лиц и выявления мошенничества<sup>30</sup>. Лица, которые подвергаются таким процедурам, задействующим ИИ, могут не знать об использовании ИИ и быть не осведомлены об этом, и у них часто нет никаких средств защиты от дискриминационного обращения со стороны государственных учреждений. Группы гражданского общества подчеркивают необходимость привлечения заинтересованных

<sup>26</sup> United Nations Educational, Scientific and Cultural Organization (UNESCO), *Steering AI and Advanced ICTs for Knowledge Societies: A Rights, Openness, Access and Multi-stakeholder Perspective* (Paris, 2019), p. 63.

<sup>27</sup> *Artificial Intelligence Index Report 2024*.

<sup>28</sup> См. <https://digitalrightsfoundation.pk/wp-content/uploads/2021/03/Policy-brief-2.pdf>; <https://www.hrw.org/report/2023/12/21/metaspoken-promises/systemic-censorship-palestine-content-instagram-and>; <https://www.apc.org/en/member/7amleh-arab-center-social-media-advancement>; <https://smex.org/>; <https://www.derechosdigitales.org/>; и <https://www.accessnow.org/>.

<sup>29</sup> См. информацию о прогностической деятельности полиции, ведущей к несоразмерно пристальному наблюдению и контролю за чернокожими общинами, по ссылке <https://naacp.org/resources/artificial-intelligence-predictive-policing-issue-brief>.

<sup>30</sup> См. <https://www.amnesty.org/en/latest/campaigns/2024/01/the-urgent-but-difficult-task-of-regulating-artificial-intelligence/>.



сторон к разработке нормативных актов в области ИИ. Было высказано предположение, что Европейский союз, несмотря на критику, продолжает использовать ИИ в дискриминационных системах надзора, таких как системы оценки рисков и прогностического анализа, чтобы форсировать практику «вытеснения»<sup>31</sup>, в то время как в контексте миграции системы ИИ разрабатываются для различных целей: а) проверка личности посредством развития технологии распознавания лиц; б) оценка рисков; в) оценка задержаний; и д) наблюдение. В таких ситуациях отмечается риск минимизации контроля со стороны человека<sup>32</sup>.

16. Учитывая, что государственные учреждения и компании, использующие ИИ для обработки данных и принятия решений, не обеспечивают прозрачность в разработке, обучении, использовании или применении ИИ, отдельные лица или группы лиц, заявляющие о нарушениях прав человека, часто не могут предоставить доказательства того, что ИИ связан с тем или иным нарушением. Нарушение может быть обусловлено в первую очередь быстрой обратной связью между различными государственными системами, связанными посредством ИИ, что приводит к жестокому обращению или принятию приносящих вред решений со стороны государственного учреждения. Преступники также используют ИИ для присвоения обманным путем денежных переводов, которые мигранты отправляют своим семьям, или денег, которые члены семьи отправляют мигрантам, как в стране происхождения, так и в стране проживания.

17. Генеральная Ассамблея признала необходимость бороться с алгоритмической дискриминацией<sup>33</sup>, выявлять факторы уязвимости, повышать доступность и предоставлять средства правовой защиты в случае нарушения прав человека<sup>34</sup>. Использование ИИ в государственных учреждениях, отвечающих за решение проблем, касающихся миграции, преступности, здравоохранения и ухода за пожилыми людьми, привлекает внимание договорных органов и специальных процедур по правам человека Организации Объединенных Наций. Комитет по ликвидации расовой дискриминации опубликовал заключительные замечания, в которых выразил обеспокоенность в связи с тем, что использование ИИ оказывает дискриминационное воздействие в контексте предоставления убежища<sup>35</sup>. Комитет по ликвидации дискриминации в отношении женщин опубликовал заключительные замечания, в которых призвал ввести надлежащие гарантии недопущения гендерных стереотипов, связанных с биометрическими системами и системами наблюдения и алгоритмического профилирования, которые правоохранительные органы используют для предотвращения и расследования преступлений, и принять меры для устранения алгоритмической предвзятости, связанной с искусственным интеллектом и алгоритмическими услугами<sup>36</sup>. Независимый эксперт по вопросу об осуществлении пожилыми людьми всех прав человека выразила признательность за предоставленную в ходе визита возможность узнать о том, какую пользу может приносить искусственный интеллект в вопросах старения, ухода и медицинского обслуживания, но при этом рекомендовала пересмотреть порядок

<sup>31</sup> Материалы, представленные Индирой Бутье, Каледонский университет Глазго.

<sup>32</sup> Там же. См. также Petra Molnar and Lex Gill, *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System* (International Human Rights Program (Faculty of Law, University of Toronto) and Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto), 2018), pp. 31–34.

<sup>33</sup> Резолюция 78/265 Генеральной Ассамблеи, пункт 6 h).

<sup>34</sup> Резолюция 78/213 Генеральной Ассамблеи.

<sup>35</sup> CERD/C/ECU/CO/23-24, пункты 45–46.

<sup>36</sup> CEDAW/C/ITA/CO/8, пункт 26. См. также аналогичные заключительные замечания в документе CEDAW/C/DEU/CO/9, пункты 27–28. Комитет выразил свою обеспокоенность по поводу ИИ и гендерного насилия в заключительных замечаниях, содержащихся в документе CEDAW/C/ESP/CO/9, пункты 21, 23, 31 и 33.

использования сбора данных, чтобы обеспечить соблюдение при этом процессе обязательств по недискриминации в отношении пожилых людей разного происхождения, то есть предлагала принять концепцию солидарности в отношении ИИ<sup>37</sup>.

18. Группы гражданского общества особо отметили негативное влияние ИИ на возможность активистов из сообщества ЛГБТКИ участвовать в акциях солидарности.

### **III. Региональные инициативы по рассмотрению институциональных процессов, связанных с искусственным интеллектом**

19. В 2024 году Организация экономического сотрудничества и развития (ОЭСР) обновила свои принципы в отношении искусственного интеллекта, особо отметив важность содействия вовлечению недопредставленных групп населения, сокращения экономического, социального, гендерного и других видов неравенства, а также поощрения разнообразия, беспристрастности, социальной справедливости, транспарентности, субъектности человека и человеческого надзора<sup>38</sup>. Проблема повышенного интереса государственных учреждений к использованию ИИ для обработки дел и принятия решений заключается в том, что многие из этих систем создаются корпорациями, а значит, существует риск отсутствия транспарентности. Государственные учреждения могут не знать об опасности возникновения таких проблем, связанных с ИИ, как «галлюцинации», когда ИИ создает материалы, не основанные на других материалах, а также о том, что при отслеживании использования информации и выявлении алгоритмической предвзятости необходимо обеспечивать контроль со стороны человека. В разработанном ОЭСР каталоге инструментов и метрик для надежного ИИ приведены примеры различных технических, образовательных и процедурных подходов к обеспечению солидарности в области ИИ<sup>39</sup>.

20. В Законе Европейского союза об искусственном интеллекте затрагиваются вопросы международной солидарности в рамках рассмотрения запрещенных видов практики, связанных с ИИ, в пункте 1) b), c) и g) статьи 5 Закона. Европейская комиссия предложила Директиву об ответственности, которую, как планируется, будут выполнять национальные суды и в которой определены стандарты доказывания<sup>40</sup>. В 2018 году Европейская комиссия приняла Европейскую этическую хартию об использовании искусственного интеллекта в судебных системах и окружающих их реалиях, в которой призвала проводить внешний аудит процесса обработки данных в целях проверки этого процесса на предмет дискриминации<sup>41</sup>. Комиссия признала, что использование ИИ при управлении миграцией, предоставлении убежища и пограничном контроле, а также при отправлении правосудия и проведении демократических процессов сопряжено с высоким риском, и поэтому в 2024 году приняла Закон об искусственном интеллекте. В Законе предусматривается создание в рамках Комиссии Европейского управления

<sup>37</sup> A/HRC/45/14/Add.1, пункт 93.

<sup>38</sup> См. <https://oecd.ai/en/ai-principles>;  
см. также <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

<sup>39</sup> См. <https://oecd.ai/en/>.

<sup>40</sup> См. <https://artificialintelligenceact.eu/the-act/>. См. также предложение о принятии директивы Европейского парламента и Совета об адаптации норм о внедоговорной гражданской ответственности к искусственному интеллекту (Директива об ответственности ИИ). Доступна по ссылке <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52022PC0496>.

<sup>41</sup> См. <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>.

по искусственному интеллекту (Управление по ИИ), которое будет контролировать эффективное исполнение и соблюдение вышеупомянутого закона поставщиками моделей ИИ общего назначения. В состав Европейского совета по искусственному интеллекту войдут представители государств-членов и корпораций. Управление по ИИ может предлагать поставщикам моделей ИИ общего назначения и соответствующим национальным компетентным органам принимать участие в создании кодексов практических мер, а гражданское, промышленное и академическое сообщества, поставщики систем ИИ с интегрированной ИИ моделью и независимые эксперты могут поддерживать этот процесс<sup>42</sup>. Этот стандарт представляется слабым, поскольку в нем нечетко обозначены возможности гражданского общества вносить непосредственный вклад в оценку воздействия.

21. Европейский суд в своем стратегическом документе определил риски дискриминации, связанные с будущим использованием ИИ, следующим образом: «Одним из основных рисков, связанных с внедрением технологии ИИ, является вероятность того, что в процессе обучения моделей ИИ может непроизвольно возникнуть предвзятость, что приведет к непреднамеренной дискриминации»<sup>43</sup>. Подкомитет по правам человека Европейского парламента провел анализ ИИ, и в результатах этого анализа был отмечен риск закрепления структурной дискриминации: «Системы ИИ могут создавать условия для увековечивания и усугубления предвзятости, что приведет к дискриминации в различных секторах, включая занятость, правоохранительную деятельность и оценку платежеспособности. Существует множество доказательств того, что ИИ может укоренять социально-экономические факторы неравенства путем воспроизведения существующих предубеждений при сборе данных, используемых для обучения ИИ, или при его разработке»<sup>44</sup>.

22. 17 мая 2024 года Комитет министров Совета Европы принял Рамочную конвенцию об искусственном интеллекте и правах человека, демократии и верховенстве права, которая будет открыта для подписания в сентябре 2024 года. Статья 10 этой рамочной конвенции, посвященная равенству и недискриминации, содержит идею разработки основанного на жизненном цикле подхода к систематическому выявлению и устранению предвзятости; в соответствии с действующим законодательством системы ИИ следует проверять на предмет дискриминации в течение всего периода их использования<sup>45</sup>.

1. Каждая Сторона принимает или поддерживает меры в целях обеспечения того, чтобы деятельность, осуществляемая в рамках жизненного цикла систем искусственного интеллекта, отвечала принципам равенства, включая гендерное равенство, и запрета дискриминации, как это предусмотрено применимым международным и внутренним правом.

2. Каждая Сторона обязуется осуществлять или поддерживать меры, направленные на преодоление неравенства для достижения непредвзятых, справедливых и равноправных результатов, в соответствии с применимыми внутренними и международными обязательствами в области прав человека в отношении деятельности в рамках жизненного цикла систем искусственного интеллекта.

<sup>42</sup> См. <https://artificialintelligenceact.eu/high-level-summary/>.

<sup>43</sup> См. [https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu\\_ai\\_strategy.pdf](https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf).

<sup>44</sup> См. [https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO\\_IDA\(2024\)754450\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf); см. также Ahmet Bilal Aytakin, “Algorithmic bias in the context of European Union anti-discrimination directives”, paper presented at the European Workshop on Algorithmic Fairness, 7–9 June 2023, Winterthur, Switzerland; и [https://fra.europa.eu/sites/default/files/fra\\_uploads/fra-2022-bias-in-algorithms\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf).

<sup>45</sup> См. <https://rm.coe.int/1680afae3c>.

23. Были и другие региональные инициативы, направленные на поощрение цифровой инклюзии, например резолюция 580 (LXXVIII) 2024 об отключении Интернета и выборах в Африке, принятая Африканской комиссией по правам человека и народов<sup>46</sup>. В ноябре 2023 года Межамериканский суд по правам человека провел диалог с судьями, в ходе которого они обсудили Глобальный инструментарий ЮНЕСКО по ИИ и верховенству права для сотрудников судебных органов. В этом инструментарии описываются проблемы, касающиеся исключения:

Проблема в том, что [...] данные могут содержать элементы предвзятости [...]. Так, например, в клинических испытаниях часто не участвуют женщины и представители «цветного» населения, что приводит к ненадлежащему представлению данных. Это может иметь серьезные последствия, если алгоритмы, обученные на основе таких данных, будут использоваться для анализа изображений кожи или определения приоритетов в уходе за пациентами. Поэтому крайне важно обеспечивать обучение алгоритмов ИИ на репрезентативных данных, чтобы не допускать подобной предвзятости и обеспечивать справедливые результаты для всех<sup>47</sup>.

24. В Инструментарии также рассматривается вопрос репрезентативности ИИ:

Цифровой разрыв во многих странах глобального Юга привел к появлению феномена «невидимость данных», что, вероятно, скажется на таких исторически маргинализированных группах, как женщины, касты, племенные общины, религиозные и языковые меньшинства, а также трудовые мигранты. Полезность и обоснованность алгоритмов ИИ, разработанных на основе легкодоступных данных, могут быть ограничены из-за предвзятости, сохраняющейся по причине невидимости данных. В связи с этим необходимо обеспечивать транспарентность и подотчетность алгоритмов.

В Инструментарии отмечается проблема косвенной дискриминации, примером которой является использование банками информации о почтовых индексах и уровне образования или дохода при оценке заявок на получение кредита, которые могут быть связаны с этническими или расовыми меньшинствами, — такой метод несет в себе риск увековечивания межсекторальных дискриминационных стратегий и практических методов. В Инструментарии говорится о том, что ИИ может помогать выявлять модели предвзятого отношения к маргинализированным группам и лицам со стороны государственных и частных субъектов, в том числе со стороны судебных органов.

25. В ходе Межамериканского диалога с судьями также обсуждался документ «Инструмент ЮНЕСКО по оценке этического воздействия ИИ», в котором содержится призыв учитывать «разнообразие проектной группы по ИИ, особенно с точки зрения гендерной и этнической принадлежности, возраста, расы, цвета кожи, происхождения, языка, религии, национальности, этнического и социального происхождения, экономического или социального положения, инвалидности и сексуальной ориентации, в том числе то, как это отражает сложность и разнообразие предполагаемой аудитории пользователей, и как это может породить предвзятость»<sup>48</sup>.

<sup>46</sup> См. <https://achpr.au.int/en/adopted-resolutions/580-internet-shutdowns-elections-africa-achpres580-lxxviii>.

<sup>47</sup> См. <https://nataliazuazo.com/2023/11/20/ai-and-the-rule-of-law-at-the-inter-american-court-of-human-rights/>. См. также <https://unesdoc.unesco.org/ark:/48223/pf0000387331>.

<sup>48</sup> См. <https://unesdoc.unesco.org/ark:/48223/pf0000386276>.

## A. Солидарность в обеспечении цифровой грамотности в целях борьбы с дезинформацией

26. Все чаще звучат призывы к регулированию «предвыборных кампаний по распространению недостоверной информации, инициированных экстремистами»<sup>49</sup>. Повышение качества материалов, генерируемых ИИ, осложняет людям задачу выявления заведомо ложной информации и недостоверных материалов, а также участились случаи отключения Интернета государствами, что оказывает несоразмерное воздействие на доступ к информации. Необходима стратегия по обеспечению международной солидарности, предусматривающая сотрудничество в правовой, технической и образовательной сферах в целях повышения уровня цифровой грамотности<sup>50</sup>. Концепция гибкой солидарности используется в качестве средства урегулирования чрезвычайных ситуаций, связанных со злонамеренным использованием ИИ, например «когда в день выборов в стране, в которой в прошлом произошел геноцид, распространяются тысячи фальшивых видеороликов со сценами насилия на этнической почве»<sup>51</sup>.

27. В 2022 году Европейский союз принял Закон о цифровых услугах, который вступил в силу в 2024 году и в котором содержится призыв к крупным онлайн-платформам и поисковым системам использовать кодексы поведения, а также принимать строго необходимые и целенаправленные меры для борьбы с дезинформацией (включая ложные дилеммы и политическую рекламу)<sup>52</sup>. Этот закон подкрепляет положения обновленного Кодекса практических мер по борьбе с дезинформацией 2022 года (применимого к саморегулируемым структурам), который включает в себя рекомендации по демонетизации деятельности по распространению заведомо ложной информации, обеспечению транспарентности политической рекламы, укреплению сотрудничества с субъектами, занимающимися проверкой фактов, и расширению доступа исследователей к данным<sup>53</sup>. Тем не менее существуют критические замечания о том, что «хотя эти события могут иметь объединяющий эффект, континентальный подход к борьбе с дезинформацией в настоящее время фрагментирован»<sup>54</sup>. Кроме того, существует опасение, что между кодексами поведения и решениями Европейского суда по правам человека могут возникнуть несоответствия<sup>55</sup>.

28. Закон об искусственном интеллекте Европейского союза также содержит положения, касающиеся ограничения влияния и воздействия «дипфейков». Согласно пункту 4 статьи 50 оператор, то есть лицо, использующее систему ИИ,

<sup>49</sup> Материалы, представленные Джейком Окечукву Эффоду.

<sup>50</sup> Там же.

<sup>51</sup> Miguel Luengo-Oroz, “Solidarity should be a core ethical principle of AI”, *Nature Machine Intelligence*, vol. 1 (November 2019). See also Patrik Hummel and Matthias Braun, “Just data? Solidarity and justice in data-driven medicine”, *Life Sciences, Society and Policy*, vol. 16, No. 8 (2020).

<sup>52</sup> Постановление (ЕС) 2022/2065 Европейского парламента и Совета от 19 октября 2022 года о едином рынке цифровых услуг и внесении изменений в Директиву 2000/31/ЕС (Закон о цифровых услугах).

<sup>53</sup> См. <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>.

<sup>54</sup> Материалы, представленные Дороти Андерсен, научным работником, занимающейся вопросами прав человека в рамках программы международного и сравнительного права, Школа права Университета им. Джорджа Вашингтона, Вашингтон, округ Колумбия.

<sup>55</sup> Paolo Cavaliere, “The Truth in Fake News: How Disinformation Laws Are Reframing the Concepts of Truth and Accuracy on Digital Platforms”, *European Convention on Human Rights Law Review*, vol. 3, No. 4 (November 2022). See also Ethan Shattock, “Fake News in Strasbourg: Electoral Disinformation and Freedom of Expression in the European Court of Human Rights”, *European Journal of Law and Technology*, vol. 13, No. 1 (2022).  
См. также <https://edmo.eu/wp-content/uploads/2022/01/Case-law-for-policy-making-Report-2022.pdf>.

должен, как правило, сообщать о том, что контент был создан с помощью ИИ. Однако в этом же законе предусмотрены исключения из этого требования о раскрытии информации, а значит, необходимо тщательно проанализировать эффективность этого закона.

29. Агентство Африканского союза по развитию призывает к созданию рамочных механизмов правовой защиты и нормативно-правовой базы для борьбы с алгоритмической дискриминацией<sup>56</sup>. Исполнительный совет Африканского союза поддерживает разработку континентальной стратегии, основанной на принятых Исполнительным советом концептуальных рамках в отношении ИИ<sup>57</sup>. В регионе наблюдается тенденция к регулированию «технологического контента», которое включает в себя цензуру или блокировку доступа к контенту с целью недопущения фальшивых новостей и/или языка ненависти, а также к регулированию «правового контента», предусматривающему обсуждение, разработку и принятие законопроектов, предписывающих правила контроля за фальшивыми новостями и языком ненависти. Существует опасность того, что оба этих вида регулирования получат слишком широкое применение или будут направлены против политической оппозиции, из-за чего выборы станут менее независимыми.

30. Саудовская Аравия проинформировала Независимого эксперта о создании Управления по вопросам данных и искусственного интеллекта Саудовской Аравии, что является значимым событием для арабской большой языковой модели (известной как АБЯМ). Эта инициатива направлена на использование технологий ИИ и цифровых приложений, чтобы поощрять культурное разнообразие и нести благо для всего человечества. Один из проектов — «Саутек» (SauTech) — направлен на локализацию технологий ИИ и сохранение местной культуры, а разработанная в рамках этого проекта система может распознавать и преобразовывать устную речь в текст для различных диалектов арабского языка, чтобы находить решения, адаптированные к местным условиям<sup>58</sup>. Управление по вопросам данных и искусственного интеллекта Саудовской Аравии повышает осведомленность об ИИ, осуществляя программы учебной подготовки. Колумбия также представила материалы с информацией об учебной подготовке по вопросам ИИ на разных уровнях<sup>59</sup>.

## **В. Гендерная солидарность в системах искусственного интеллекта**

31. Кроме того, все больше внимания уделяется тому, что подход, основанный на принципе международной солидарности, может способствовать борьбе с гендерной дискриминацией в отношении ИИ<sup>60</sup>. В докладе «Индекс данных, касающихся искусственного интеллекта» 2024 года говорится о том, что в европейских странах мужчин, получивших высшее образование в области информационных технологий, больше, чем женщин, при этом этот гендерный разрыв сокращается медленными темпами<sup>61</sup>. Комитет по ликвидации дискриминации в отношении женщин опубликовал заключительные замечания к ряду докладов государств-

<sup>56</sup> См. <https://www.nepad.org/publication/ai-and-future-of-work-africa-white-paper>.

<sup>57</sup> См. <https://au.int/en/newsevents/20240419/multistakeholder-consultative-sessions-development-continental-strategy>.

<sup>58</sup> Материалы, представленные Саудовской Аравией.

<sup>59</sup> Материалы, представленные Колумбией.

<sup>60</sup> Keratso Georgiadou, “Solidarity, Gender, Dialogue in the AI era”, in *Kritische Pädagogik und Bildungsforschung: Anschlüsse an Paulo Freire*, Wassilios Baros, Rita Braches-Chyrek, Solvejg Jobst and Joachim Schroeder, eds. (Wiesbaden, Germany, Springer Fachmedien, 2024), pp. 477–489.

<sup>61</sup> *Artificial Intelligence Index Report 2024*.



участников, в которых выразил свою обеспокоенность по поводу недостаточной представленности девочек и женщин в сфере искусственного интеллекта и по поводу отсутствия конкретных мер по предотвращению интеграции гендерных стереотипов и алгоритмической гендерной дискриминации и насилия при написании кодов для ИИ и призвал к созданию эффективных гарантий<sup>62</sup>.

32. В своих материалах Саудовская Аравия отметила, что женщины составляют более 50 процентов технического персонала Управления по вопросам данных и искусственного интеллекта Саудовской Аравии<sup>63</sup>. Созданный Саудовской Аравией Международный центр исследований и этики искусственного интеллекта в сотрудничестве с компанией «Гугл клауд» приступил к реализации инициативы “Elevate” («Повышение»), направленной на учебную подготовку 25 000 женщин в области данных и ИИ. Институт демократии и прав человека в Перу сообщил, что президент Перу поддержала программу обучения для предприятий, разрабатывающих проекты с использованием ИИ. Граждане, которые зарегистрировались на программу с мая 2024 года, имели возможность подать заявку на получение стипендий на обучение, 50 процентов из которых были предоставлены женщинам, находящимся в уязвимом положении<sup>64</sup>.

33. Ресурсный центр по вопросам бизнеса и прав человека описывает существующие проблемы, связанные с тем, что ИИ создает условия, противоречащие проявлению солидарности в отношении женщин, следующим образом<sup>65</sup>:

Женщины и гендерно неконформные лица продолжают испытывать несоизмерно сильное воздействие плохо продуманных, плохо разработанных и плохо внедренных технологий, о чем свидетельствует то, что новейшие технологические инструменты, включая генеративный ИИ, выпускаются с недостатками и элементами предвзятости, которые усиливают сексизм и гендерные стереотипы, создают питательную среду для социального контроля над женщинами, лишают женщин-пациентов прав, являются дискриминационными по отношению к женщинам-соискателям, вносят вклад в преследование правозащитниц и ставят женщин в невыгодное положение при получении социальных пособий, а также наносят другой вред.

34. Женщины все чаще слышат в свой адрес порожденные ИИ ненавистнические высказывания, получают соответствующие материалы сексуального характера без согласия и становятся жертвами кибербуллинга, размещаемых из мести порнографических материалов и stalkingа, что отрицательно сказывается на их способности реализовывать свои личные и профессиональные устремления. Организации гражданского общества рассказали о новаторской инициативе, направленной на прекращение гендерного насилия и включающей в себя учебный курс по вопросам цифровой безопасности, организованный в Международном университете Кампалы в 2023 году<sup>66</sup>. Они также отметили, что использование социальных сетей играет важную роль в проявлении международным сообществом солидарности в отношении требования о привлечении к уголовной ответственности за фемцид и проведении законодательной реформы в Кении,

<sup>62</sup> См. CEDAW/C/ITA/CO/8, CEDAW/C/DEU/CO/9, пункты 27 и 28; CEDAW/C/ESP/CO/9, пункты 21 и 23, 31 и 33; CEDAW/C/TJK/CO/7, пункты 47, 48 и 55. См. также заключительные замечания, содержащиеся в документе CEDAW/C/TUR/CO/8, пункт 46; CEDAW/C/CRI/CO/8, пункты 27, 28 и 37.

<sup>63</sup> Материалы, представленные Саудовской Аравией.

<sup>64</sup> Материалы, представленные Институтом демократии и прав человека, Перу.

<sup>65</sup> См. [https://media.business-humanrights.org/media/documents/BHRR\\_C\\_Submission\\_\\_Gender\\_and\\_Role\\_of\\_Business\\_2023\\_j2HkLgP.pdf](https://media.business-humanrights.org/media/documents/BHRR_C_Submission__Gender_and_Role_of_Business_2023_j2HkLgP.pdf).

<sup>66</sup> Материалы, представленные консорциумом «Наши голоса, наше будущее».



объявления положения дел, касающегося фемицида и насилия в отношении женщин, национальным бедствием, а также создания комиссии, которой будет поручено бороться с этими преступлениями и прилагать усилия, чтобы разорвать этот порочный круг безнаказанности<sup>67</sup>. Организации гражданского общества призвали к расширению международного сотрудничества между технологическими компаниями, неправительственными организациями (НПО) и правительствами для разработки эффективных и этичных инструментов модерации контента на основе ИИ, позволяющих выявлять и фиксировать случаи насилия в отношении женщин в Интернете и адаптируемых к различным языкам и культурным контекстам.

### **С. Соблюдение принципа солидарности при разработке искусственного интеллекта**

35. Во второй пересмотренной редакции Глобального цифрового договора заявлено, что он будет осуществляться в духе глобальной солидарности, и особо отмечается принцип инклюзивности, поскольку уделяется особое внимание тому, что компании, занимающиеся цифровыми технологиями, и разработчики таких технологий должны взаимодействовать с обладающими разными способностями пользователями из всех слоев общества, чтобы интегрировать их взгляды и потребности в жизненный цикл цифровых технологий, добиваться справедливого распределения цифровых благ и поддерживать доступность цифровых технологий и языковое и культурное разнообразие в цифровой среде<sup>68</sup>. Кроме того, международное сообщество призывает учитывать при разработке и реализации национальных и местных стратегий обеспечения цифровой связности потребности лиц, находящихся в уязвимом положении, и жителей недостаточно обслуживаемых, сельских и отдаленных районов, а также подчеркивает необходимость оказания целенаправленных и специальных услуг по наращиванию потенциала для женщин и девочек, детей и молодежи, а также пожилых людей, людей с инвалидностью и лиц, находящихся в уязвимом положении, и обеспечения конструктивного участия таких групп людей в создании и внедрении программ. Кроме того, международное сообщество планирует подготавливать и проводить национальные обследования цифровой инклюзивности с систематической разбивкой данных по возрасту, инвалидности и гендерной принадлежности, чтобы выявлять пробелы в обучении и определять приоритеты в конкретных условиях.

36. Глобальный цифровой договор поддерживает принятие национального законодательства в области ИИ и эффективных механизмов надзора и правовой защиты. В нем также содержится призыв к технологическим компаниям и разработчикам ИИ признавать обязательство соблюдать права человека и осуществлять процедуры должной осмотрительности и оценки воздействия. В этом договоре отмечается, что технологическим компаниям и разработчикам ИИ следует разработать в консультации с правительствами и другими заинтересованными сторонами отраслевые механизмы подотчетности, способствующие повышению прозрачности их систем и процессов, определить обязанности и принять на себя обязательство соблюдать стандарты и подготавливать публичные доклады, подлежащие аудиту. В Договоре к платформам социальных сетей обращен призыв создавать безопасные, надежные и доступные механизмы подачи жалоб, с помощью которых пользователи и их защитники могли бы сообщать о нарушениях политики (включая специальные механизмы подачи жалоб, адаптированные для

<sup>67</sup> Там же.

<sup>68</sup> Доступно по ссылке <https://www.un.org/techenvoy/global-digital-compact>.

детей и лиц с инвалидностью). В нем отмечается, что технологическим компаниям, работающим в сфере ИИ, и платформам социальных сетей следует повышать прозрачность и подотчетность своих систем, включая условия предоставления услуг, политику модерации контента и алгоритмы, а также стратегии обработки личных данных пользователей на местных языках, чтобы дать пользователям возможность делать осознанный выбор и давать или отзываться информированное согласие.

37. ЮНЕСКО подчеркивает, что обеспечение инклюзивности при разработке восприимчивых многосторонних механизмов ИИ является одним из элементов прозрачности<sup>69</sup>:

- Поощряет ли правительство участие других заинтересованных сторон в управлении ИИ на национальном уровне?
- Существуют ли в стране действующие ассоциации специалистов, потребителей и других заинтересованных сторон в области ИИ?
- Занимается ли правительство активным вовлечением других групп заинтересованных сторон в процесс разработки политики глобального управления ИИ?

ЮНЕСКО призывает внедрять индикаторы для обеспечения инклюзивности путем определения того, насколько различные группы заинтересованных сторон (женщины и гендерно вариативные лица, молодежь и маргинализированные группы) вовлечены в разработку политики в области ИИ и насколько активно они участвуют в международных и региональных форумах, процессах и механизмах, связанных с ИИ.

38. Было высказано следующее предположение: «технологии ИИ в социально-технологической системе могут способствовать социальным взаимодействиям или формировать их и, следовательно, могут либо подрывать солидарность, либо укреплять ее»<sup>70</sup>. Чрезвычайно важно привлекать заинтересованные стороны к принятию решений о разработке ИИ, учитывая, что их значимость может меняться со временем, а значит, возможны непредвиденные проблемы с точки зрения солидарности. Права человека — это мерило солидарности в отношении ИИ<sup>71</sup>. Считается, что использование ИИ может привести к таким ситуациям, в которых нарушаются права на неприкосновенность частной жизни из-за применения технологии распознавания лиц, или нарушается принцип недискриминации женщин или меньшинств при приеме на работу, или людей лишают возможности достигать своих жизненных целей (или реализовывать дело всей жизни), например отказывают в кредите на приобретение жилья или получение образования. Оценкам воздействия на права человека следует отводить центральное место при разработке ИИ с учетом принципа солидарности. Более того, следует определять интересы и потребности пользователей и учитывать их в процессе разработки. В условиях коллективной разработки пользователи смогут обсуждать варианты технического исполнения и участвовать в процессе принятия решений, который должен включать коллективное принятие решений, способное обеспечить солидарность в рамках группы. Однако сосредоточение внимания только на пользователях не позволяет учитывать косвенные заинтересованные стороны, например тех, кто подвергается непреднамеренному воздействию.

<sup>69</sup> ЮНЕСКО, «Управление развитием и использованием искусственного интеллекта и передовых информационно-коммуникационных технологий в интересах обществ, основанных на знаниях».

<sup>70</sup> Catharina Rudschies, “Exploring the Concept of Solidarity in the Context of AI: An Ethics in Design Approach”, *Digital Society*, vol. 2, No. 1 (2023).

<sup>71</sup> Там же, на странице 12.

Стратегия разработки ИИ с учетом принципа солидарности требует уделения внимания не только пользователям, но и всем затрагиваемым заинтересованным сторонам в социально-технологической системе. Это подразумевает не только оценку вреда и пользы в целом, но и оценку того, как вред и польза распределяются в обществе и между группами заинтересованных сторон. Основным вопросом заключается в том, как заставить технологические компании принять на себя обязательство проявлять солидарность и разделять риски в дополнение к совместному использованию выгод. Подход к разработке ИИ с учетом принципа солидарности позволит проводить оценку рисков в области прав человека, определять право людей на самореализацию и регулировать распределение ответственности за причиненный вред, а также распределение выгод от обработки данных и принятия решений<sup>72</sup>. Следовательно, потребуется также политическая солидарность, чтобы обеспечить поддержку коллективного принятия решений, а также основанная на принципе солидарности система управления ИИ для урегулирования последствий использования ИИ.

39. Подход к разработке ИИ с учетом принципа солидарности может потребовать формулирования правил в различных областях, таких как здравоохранение или энергетика. Диалог в рамках разработки и развития ИИ должен поощрять сотрудничество между пользователями, разработчиками технологий ИИ и другими заинтересованными сторонами в целях поиска решений в области ИИ. Необходимо создавать пути взаимодействия для поддержки диалога с разработчиками сценариев использования ИИ, разработчиками решений на основе ИИ или любыми другими разработчиками применимых технологий ИИ/машинного обучения, чтобы обеспечивать постоянное информирование для рассмотрения меняющихся рисков и преимуществ, связанных с использованием решений на основе ИИ. Также необходимо разрабатывать специальные планы коммуникаций и взаимодействия с учетом рисков, чтобы клиенты могли легко понимать разъяснения о том, как было разработано определенное решение ИИ, каковы его эксплуатационные характеристики и как осуществляется его обслуживание, а также как это решение увязывается с передовой практикой и соответствует нормативным требованиям.

40. Признание и уважение должны считаться элементами основанного на принципе солидарности управления данными, которое обеспечивает «концептуальное место для конкретных случаев пережитого опыта маргинализации и чувства несправедливости в контекстах, требующих больших объемов данных»<sup>73</sup>. Солидарность как принцип разработки ИИ должна подразумевать устойчивое справедливое воздействие в целях всеобщего процветания, распределения бремени и предотвращения неравенства<sup>74</sup>. Ученые предлагают создать всеобъемлющую систему разработки ИИ с учетом принципа солидарности, основанную на стимулах к предотвращению неравенства, таких как выплата отчислений за каждый случай использования системы ИИ, обученной на чьих-то данных, и премий врачам, применяющим модель ИИ для диагностики, или вознаграждение лиц, создающих текст для автоматического генератора текстов на основе ИИ, за каждую публичную статью, написанную роботом<sup>75</sup>. Другие призывают к подходу, который основан на принципе солидарности и предусматривает измерение внешнего воздействия на общество путем пропаганды гарантий индивидуального контроля над обрабатываемыми данными<sup>76</sup>. Еще одна точка зрения заключается в том, что необходимо создать систему управления данными на основе

<sup>72</sup> Там же, на странице 13.

<sup>73</sup> Hummel and Braun, “Just data? Solidarity and justice in data-driven medicine”.

<sup>74</sup> Luengo-Oroz, “Solidarity should be a core ethical principle of AI”.

<sup>75</sup> Ibid.

<sup>76</sup> Hummel and Braun, “Just data? Solidarity and justice in data-driven medicine”.

принципа солидарности, позволяющую обеспечивать наглядность процессов обработки данных на благо общества<sup>77</sup>. Сторонники такой точки зрения полагают, что это может побудить корпоративные и государственные заинтересованные стороны в области данных разделять риски и преимущества получения доступа к данным, их подготовки и обмена ими. Управление данными на основе принципа солидарности представляется как метод обеспечения наглядности процессов, касающихся подготовки набора данных, с целью выявления и устранения дискриминации по признаку принадлежности к маргинализованной группе<sup>78</sup>. Стороны, придерживающиеся вышеизложенной точки зрения, призывают к коллективным действиям, основанным на использовании солидарности как одного из принципов управления данными, направленного на создание общедоступных наборов данных, чтобы начать укреплять доверие и повышать уровень подотчетности. Существует мнение, согласно которому международное сообщество переживает переломный момент: управление данными на основе принципа солидарности должно способствовать созданию инфраструктуры и должно быть организовано в соответствии с демократическими ценностями, такими как инклюзивность<sup>79</sup>. В свою очередь, возможно, потребуется создать независимый орган по вопросам глобального управления, в состав которого войдут представители промышленности, государств, гражданского общества, международных организаций и академических кругов и который будет продвигать основанные на правах человека правила в отношении ИИ<sup>80</sup>. Можно рассмотреть подготовленную компанией «Антропик» «Конституцию Клода», в которой изложены принципы, стимулирующие, помимо соблюдения прав человека, учет точек зрения незападных культур<sup>81</sup>.

41. Некоторые технологические компании представили материалы, в которых описаны элементы, необходимые для внедрения систем ИИ, по умолчанию учитывающих принцип международной солидарности:

- a) проведение учебных занятий, разработанных с учетом культурных особенностей и с соблюдением принципов доступности и инклюзивности;
- b) обучение алгоритмов на разнообразных наборах данных, чтобы они представляли людей, которых они обслуживают;
- c) регулярный аудит, тестирование по различным демографическим группам и применение методов машинного обучения, основанных на принципе справедливости;
- d) определение этических принципов ИИ, в которых при разработке и внедрении алгоритмов приоритетное внимание уделяется соблюдению принципов справедливости, транспарентности и подотчетности;
- e) постоянный мониторинг и корректировка систем ИИ в целях предотвращения дискриминации;
- f) взаимодействие с правозащитными группами, НПО и представителями уязвимых групп населения для выявления аспектов, вызывающих обеспокоенность, и сбора мнений о влиянии технологий;

<sup>77</sup> Mercedes Bunz and Photini Vrikki, "From Big to Democratic Data: Why the Rise of AI Needs Data Solidarity", in *Democratic Frontiers: Algorithms and Society*, Michael Filimowicz, ed. (London, Taylor & Francis, 2022).

<sup>78</sup> Ibid.

<sup>79</sup> Ibid.

<sup>80</sup> Ana Beduschi, "Human rights and the governance of artificial intelligence", research brief (Geneva Academy of International Humanitarian Law and Human Rights, 2020).

<sup>81</sup> См. <https://www.anthropic.com/news/claude-constitution>.

- g) постоянное обучение команд по вопросам алгоритмической справедливости и этических последствий использования ИИ;
- h) разработка стратегий модерации контента, запрещающих распространение ложной информации, ненавистнических высказываний и вредного контента;
- i) обеспечение соблюдения правил с помощью автоматизированных инструментов и сотрудников-модераторов, обученных выявлять и удалять заведомо ложную информацию;
- j) сотрудничество с надежными организациями по проверке фактов;
- k) сосредоточение внимания на обеспечении прозрачности алгоритмов с целью ограничения непреднамеренного распространения заведомо ложной информации;
- l) разработка алгоритмов, направленных на продвижение достоверных источников и уменьшение доступности недостоверного контента;
- m) проведение информационно-просветительских кампаний и размещение уведомлений на платформах в целях обучения пользователей тому, как выявлять заведомо ложную информацию и как сообщать о ней;
- n) сотрудничество с правительственными и неправительственными организациями и научными учреждениями для обмена опытом и передовыми методами эффективной борьбы с дезинформацией.

#### IV. Нацеленность стран на положительную практику

42. Испания сообщила о том, что разработала национальную стратегию в области искусственного интеллекта и учредила орган по контролю за ее реализацией<sup>82</sup>. Испания заявила, что создаст этическую и нормативную базу, которая укрепит защиту индивидуальных и коллективных прав, чтобы гарантировать инклюзивность и социальное благополучие. Она планирует разработать хартию цифровых прав и создать модель этических аспектов управления ИИ на национальном уровне через консультативный совет по вопросам ИИ в сотрудничестве с консультативным советом по цифровым преобразованиям. Статья 23 закона № 15/2022 от 12 июля 2022 года гласит, что необходимо принимать меры для снижения предвзятости и в то же время способствовать повышению прозрачности и подотчетности при использовании ИИ в решениях государственных органов. К этим мерам относятся анализ используемых для обучения данных и процесса разработки и их оценка на предмет дискриминационного влияния<sup>83</sup>. Кроме того, Испания стремится создать форумы для диалога, повышения осведомленности и национального и международного участия в области ИИ, чтобы способствовать налаживанию связей между правительством, научным сообществом, частным сектором и гражданским обществом.

43. Саудовская Аравия сообщила, что у ее Управления по вопросам данных и искусственного интеллекта есть система для организации, подготовки и обработки данных ИИ и предоставления государственных услуг, с тем чтобы надежным образом способствовать цифровым преобразованиям и сохранению данных путем привлечения заинтересованных сторон к разработке и внедрению ИИ<sup>84</sup>.

<sup>82</sup> См. <https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/ficheros/National-Strategy-on-AI.pdf>.

<sup>83</sup> Материалы, представленные Испанией.

<sup>84</sup> Материалы, представленные Саудовской Аравией.

Управление продвигает семь этических принципов ИИ, к ним относятся: справедливость; конфиденциальность и защита; гуманность; обеспечение социальных и экологических выгод; надежность и безопасность; транспарентность и объяснимость; и подотчетность и ответственность<sup>85</sup>. Управление объявило о программе стимулирования, призванной помочь компаниям добровольно соблюдать этические принципы ИИ. Процесс начинается с выявления и оценки всех потенциальных рисков и степени их воздействия.

44. Малайзия представила материалы, в которых описывается предлагаемый ею механизм надзора — структура по координации и внедрению ИИ, которая будет выступать в качестве правительственного органа, занимающегося всеми связанными с ИИ вопросами<sup>86</sup>. Малайзия планирует создать комитет по прогнозированию, который будет заниматься изучением общей обстановки, прогнозированием и информационно-пропагандистской деятельностью по вопросам политики.

45. Германия использует ИИ в рамках проекта Match'In для более эффективной интеграции беженцев в общины, проводя с беженцами консультации с целью выявления их предпочтений в отношении, в частности, жилья, работы/профессионального опыта, хобби, отдыха, медицинского обслуживания, семейного положения и религиозности общины<sup>87</sup>. Такой подход способствует признанию права беженцев на самореализацию в соответствии с принципом международной солидарности. Отмечается, что Match'In — это единственный проект, в рамках которого при разработке программ учитывается точка зрения беженцев<sup>88</sup>.

46. Доминиканская Республика сообщила, что у нее есть национальная стратегия в области искусственного интеллекта, приоритетное внимание в которой уделяется использованию ИИ в государственном управлении в таких ключевых секторах, как правосудие, здравоохранение, образование, защита окружающей среды и безопасность, а также использованию предиктивных аналитических моделей для разработки услуг для граждан. Эта страна планирует разработать кодекс, в котором будут предусмотрены превентивные меры, процедурные гарантии и механизмы подотчетности для обеспечения ответственного внедрения ИИ. Кроме того, будут созданы механизмы надзора для контроля за соблюдением этических норм. Деятельность Центра по развитию кадрового потенциала и инновациям YoSoyFuturoRD будет ориентирована на уязвимые сектора. Доминиканская Республика предлагает создать стойкую и основанную на совместных усилиях региональную экосистему ИИ, которая будет способствовать технологическому прогрессу, экономическому и социальному развитию и сотрудничеству в регионе. Страна заявляет, что создаст механизмы надзора, компенсации и обжалования для устранения ущерба, нанесенного системами ИИ, с тем чтобы обеспечить защиту прав своих граждан.

47. Муниципалитет Росон в Аргентине использует ИИ для продвижения программ социальной помощи и стимулирования, таких как программа «Сеть социальной и солидарной экономики», для создания устойчивых продовольственных схем, основанных на принципах продовольственного суверенитета. Кроме того, при муниципалитете Росона функционируют районные центры обслуживания («Мунисерка»), которые занимаются обработкой запросов на благоустройство общественных пространств и улиц, возбуждением дел и консультированием по соответствующим вопросам, а также приемом заявок на муниципальном уровне.

<sup>85</sup> В представленных материалах Колумбия также отмечает соблюдение таких принципов, как транспарентность, конфиденциальность, контроль со стороны человека и недискриминация, что помогает обеспечивать инклюзивность и справедливость технологий.

<sup>86</sup> Материалы, представленные Малайзией.

<sup>87</sup> См. <https://matchin-projekt.de/en/>.

<sup>88</sup> См. [https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum\\_afar\\_9-1-23.pdf](https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf).

## **А. Солидарность в сфере здравоохранения и искусственный интеллект**

48. Солидарность в сфере здравоохранения должна лежать в основе правил использования ИИ для обеспечения доступа и справедливости, однако достижение этой цели усложняется из-за приватизации сферы здравоохранения во многих странах<sup>89</sup>. Существующее неравенство в области здравоохранения, обусловленное, в частности, такими факторами, как расовая и классовая принадлежность и возраст, может быть воспроизведено в медицинских системах на основе ИИ, при разработке которых не уделялось внимание проблеме структурной несправедливости. Необходимо, чтобы используемый для оказания медицинской помощи ИИ создавался с учетом особенностей общинных систем здравоохранения и, в частности, применялся в районах с недостатком ресурсов, способствуя охвату медицинским обслуживанием населения, живущего в условиях недостатка ресурсов. Некоторые выдвигают идею о том, что можно создать ИИ в сфере здравоохранения, который будет основан на принципе солидарности и будет предусматривать для людей возможность предоставления и удаления данных, путем контроля потоков данных, создания инфраструктуры, позволяющей обеспечивать контроль, и сосредоточения внимания на том, как данные будут использоваться в управлении, чтобы предотвращать и уменьшать несправедливость<sup>90</sup>. В Бразилии действует программа семейных пособий (Bolsa Família), в рамках которой цифровые инструменты используются для распределения помощи семьям с низким уровнем дохода. Эта программа предусматривает осуществление денежных переводов только в случае выполнения определенных условий, например посещения школы или прохождения медицинского осмотра, и тем самым способствует сокращению масштабов нищеты и развитию человеческого капитала. Малайзия использует ИИ для медицинской диагностики, подбора индивидуального лечения, предоставления информации через окна чатов и проведения прогностического анализа для определения упреждающих мер<sup>91</sup>.

49. Саудовская Аравия представила информацию о финансировании деятельности, связанной с разработкой ИИ в сфере здравоохранения на основе солидарности. В секторе здравоохранения этой страны действует Центр передового опыта в области искусственного интеллекта. Саудовские эксперты и инженеры разработали решение Eupai, призванное произвести революцию в диагностической медицине на региональном уровне. Диабетическая ретинопатия — это заболевание, вызываемое диабетом 1 и 2 типов; она служит одной из основных причин слепоты в Саудовской Аравии. Для предотвращения последующих осложнений жизненно важна ранняя диагностика, проведению которой способствует решение Eupai, обеспечивая точное выявление и правильный диагноз. В 2022 году приложение Tawakkalna, разработанное Управлением по вопросам данных и искусственного интеллекта Саудовской Аравии, получило Премию Организации Объединенных Наций за вклад в развитие государственной службы за то, что позволяло принимать инновационные меры реагирования на пандемию коронавирусного заболевания (COVID-19).

<sup>89</sup> Nicolas Terry, “Of Regulating Healthcare AI and Robots”, *Yale Journal of Law and Technology*, special issue No. 133 (2019).

<sup>90</sup> Hummel and Braun, “Just data? Solidarity and justice in data-driven medicine”.

<sup>91</sup> Материалы, представленные Малайзией.



## V. Солидарность трудящихся и искусственный интеллект

50. Считается, что фрагментация рынка труда, вызванная использованием ИИ, мешает работникам организовываться в духе солидарности: «Легкость, с которой компания-платформа может нанимать новых работников, иногда даже со всего мира, порождает конкуренцию и изоляцию и значительно препятствует формированию коллективной идентичности. Одним словом, работа на платформах приводит к фрагментации коллективной идентичности работников и подрывает коллективные действия, особенно действия на основе солидарности и доверия»<sup>92</sup>.

51. Тем не менее существует мнение, что программы-посредники, призванные обеспечивать солидарность, могут помогать работникам в кризисные периоды<sup>93</sup>. Один из центральных вопросов касается того, можно ли переосмыслить ИИ как инструмент укрепления солидарности трудящихся<sup>94</sup>. Банк социального обеспечения Уругвая создал централизованные, автоматизированные процедуры для работодателей, способствующие осуществлению самоуправления через онлайн-сервисы, мобильное приложение и систему многоканальной помощи. Согласно сообщениям чат-бот ответил на 97 процентов запросов со 100-процентным показателем удовлетворенности. В результате удалось добиться значительного, на 24,4 процента, снижения уровня уклонения от уплаты взносов на социальное страхование: 57 процентов работодателей зарегистрировались в онлайн-сервисе, а 42 процента осуществляли платежи в режиме онлайн<sup>95</sup>. Главное управление Аргентины по вопросам профессиональных рисков успешно внедрило чат-бота на основе ИИ под названием Julieta<sup>96</sup>. С одной стороны, появляются новые формы горизонтальной солидарности, которые отличаются большей открытостью и разнообразием. Существует также тенденция к «распределенному дискурсу», «при котором бюрократические барьеры между активистами и представителями профсоюзов устранены, транспарентность повышается, а рядовые члены получают больше возможностей для того, чтобы бросать вызов олигархическим профсоюзным структурам и переформировать их»<sup>97</sup>. Более того, развивается такое явление, как «ускоренный плюрализм», под которым понимается «продолжающаяся фрагментация групповой политики, основанной на интересах, в пользу постоянно меняющейся групповой политики, основанной на проблемах, и появление новых безопасных пространств для демократических дискуссий среди женщин и других работников, которые ранее находились в маргинализированном положении из-за традиционных профсоюзных структур»<sup>98</sup>. Примером разработки ИИ с учетом солидарности можно считать расширение возможностей женщин, занятых в сельском хозяйстве, получать доступ к ИИ,

<sup>92</sup> Tammy Katsabian, “The rule of technology – how technology is used to disturb basic labor law protections”, *Lewis and Clark Law Review*, vol. 25, No. 3 (2021).

<sup>93</sup> Saiph Savage and Mohammad H. Jarrahi, “Solidarity and AI for transitioning to crowd work during COVID-19”, paper produced for virtual symposium on The New Future of Work, August 2020. See also Kurt April, “AI-Induced Solidarity Economy: The Need for Stewardship Orientation”.

<sup>94</sup> Frances Flanagan and Michael Walker, “How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia”, *New Technology, Work and Employment*, vol. 36, No. 2 (2021).

<sup>95</sup> См. <https://www.issa.int/sites/default/files/documents/2024-06/2-AI%20in%20SecSoc%202024.pdf>.

<sup>96</sup> Ibid.

<sup>97</sup> См. Andy Hodder, David Houghton, “Union use of social media: a study of the University and College Union on Twitter”, *New Technology, Work and Employment*, vol. 30, No. 3 (November 2015).

<sup>98</sup> См. Anne-Marie Greene and Gill Kirton, “Possibilities for remote participation in trade unions: mobilising women activists”, *Industrial Relations Journal*, vol. 34, No. 4 (October 2003).

способному прогнозировать урожайность и погодные условия<sup>99</sup>. Еще одним примером является движение против апартеида, возглавляемое работниками технологической сферы<sup>100</sup>. С другой стороны, Калифорнийская ассоциация медсестер и медбратьев активно требует разработки гарантий в отношении ИИ для защиты пациентов от некачественного лечения<sup>101</sup>. Тем не менее было отмечено, что «работодатели со своей стороны также мобилизовали усилия в онлайн-пространстве, а неопределенный правовой статус высказываний в социальных сетях дополнительно снижает значимость высказываний сотрудников в Интернете»<sup>102</sup>.

52. Таким образом, контекст и тип профсоюзной деятельности, а также политические и организационные различия рабочих движений будут влиять на использование ИИ в привязке к принципу международной солидарности в вопросах борьбы за власть, отражая групповую идентичность, внутреннюю сплоченность, культуру, стратегию, управление и сообщество<sup>103</sup>.

## V. Процедуры рассмотрения жалоб, касающихся вопросов должной осмотрительности

53. Все большее беспокойство по поводу прозрачности приводит к росту числа требований о пересмотре чрезмерных полномочий корпораций в отношении доступа к персональным данным и их использования. Так, например, компания «Мета» объявила о том, что планирует внедрить настройку по умолчанию, в соответствии с которой контент пользователей будет использоваться для обучения моделей искусственного интеллекта. Норвежский совет потребителей посчитал процесс отключения такой настройки сложным и подал в суд жалобу на эту компанию за нарушение Общего регламента Европейского союза по защите данных<sup>104</sup>. В Руководящих принципах предпринимательской деятельности в аспекте прав человека в принципе 15 сформулировано обязательство компаний проявлять должную осмотрительность и содержится призыв к применению превентивных подходов, а также к осуществлению процесса правовой защиты; в принципе 25 заложены основы для доступа к правосудию; а в принципе 29 к предприятиям обращен призыв создавать механизмы правовой защиты.

54. В Глобальных принципах добросовестности в информационной среде Организации Объединенных Наций содержится призыв к созданию надзорных механизмов и проведению регулярных внешних и независимых аудиторских проверок соблюдения прав человека, которые бы охватывали, в частности, такие аспекты, как условия предоставления услуг, стандарты сообщества, рекламные стратегии, модерация контента, процессы рассмотрения жалоб, предоставление исследователям доступа к данным и оценка воздействия на положение в отношении уязвимости и маргинализации, гендерного равенства и прав детей.

<sup>99</sup> См. CEDAW/C/OMN/CO/4. Информацию об искусственном интеллекте и прогнозировании погоды см. по ссылке <https://deepmind.google/discover/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/>.

<sup>100</sup> См. <https://www.notechforapartheid.com/>.

<sup>101</sup> См. <https://www.nationalnursesunited.org/press/cna-demand-patient-safeguards-against-artificial-intelligence>

<sup>102</sup> Flanagan and Walker, “How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia”. See also Louise Thornthwaite, “Chilling times: Social media policies, labour law and employment relations”, *Asia Pacific Journal of Human Resources*, vol. 54, No. 3 (August 2015).

<sup>103</sup> Ibid.

<sup>104</sup> См. <https://www.forbrukerradet.no/side/legal-complaint-against-metas-use-of-personal-content-for-ai-training/>.

Аудиторские проверки должны быть публичными, доступными и понятными для всех пользователей<sup>105</sup>.

55. В своей Рекомендации об этических аспектах искусственного интеллекта ЮНЕСКО отметила, что государства должны создавать механизмы надзора, которые будут обеспечивать участие всех заинтересованных сторон, и поощрять все заинтересованные стороны развивать права человека. В Директиве (ЕС) 2024/1760 Европейского парламента и Совета от 13 июня 2024 года о должной осмотрительности в области корпоративной устойчивости в статье 14 закреплено обязательство государств обеспечивать наличие у компаний справедливых, общедоступных, предсказуемых и транспарентных механизмов подачи жалоб для лиц, организаций гражданского общества, правозащитников, профсоюзов и других сторон, чтобы более эффективно бороться с нарушениями прав человека и экологическими последствиями деятельности компаний, их дочерних предприятий или других субъектов в цепочке деятельности<sup>106</sup>. Необходимо, чтобы эти меры отвечали стандартам конфиденциальности и не допускали преследования физических или юридических лиц, подавших жалобу. Вышеупомянутая директива создает основу, дополняющую механизм коммуникации, о создании которого говорится в пересмотренном проекте декларации о праве на международную солидарность (пункт 3 статьи 6). Кроме того, в этом пересмотренном проекте декларации (в пункте 3 статьи 8) определены гибкие рамки для возможного создания стратегий и практических методов, касающихся разработки ИИ на основе принципа международной солидарности, и для обмена соответствующей информацией в ходе универсального периодического обзора. Согласно пункту 3 статьи 8 пересмотренного проекта декларации государства обязаны принимать меры в рамках своих соответствующих возможностей для содействия защите реального и виртуального пространства для коммуникации, включая доступ к Интернету и инфраструктуре, с тем чтобы отдельные лица и народы могли обмениваться идеями о концепции солидарности. Эти нормы служат основой для поддержки творческих механизмов надзора за разработкой ИИ с учетом принципа международной солидарности, которые могут способствовать обеспечению более широкого вовлечения гражданского общества.

## VI. Вывод

56. Независимый эксперт согласна с мнением о том, что необходимо создать глобальную модель управления, предусматривающую участие многих заинтересованных сторон, для предотвращения и устранения дискриминации и других нарушений прав человека в системах ИИ<sup>107</sup>. Что касается управления, то Консультативный орган высокого уровня по искусственному интеллекту рекомендует использовать взаимодополняемый подход, охватывающий государственных и частных субъектов и распространяющийся на все юрисдикции, включая международный, региональный и национальный уровни, например Закон об искусственном интеллекте Европейского союза<sup>108</sup>. Тем не менее ключевым признаком развивающейся системы является использование добровольных кодексов и соглашений, таких как соглашение о борьбе с дипфейками, направленное на противодействие вмешательству в выборы и подписанное технологическими

<sup>105</sup> См. <https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf>.

<sup>106</sup> См. [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L\\_202401760](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401760).

<sup>107</sup> См. [https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality?gad\\_source=CjwKCAjw4f6zBhBVEiwATEHFVvzWNYAYvJV56epBISUMbQTVZ4hm\\_tCs\\_n\\_VJGGvIzoRMMXfutYebdhoCzqEQAvD\\_BwE](https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality?gad_source=CjwKCAjw4f6zBhBVEiwATEHFVvzWNYAYvJV56epBISUMbQTVZ4hm_tCs_n_VJGGvIzoRMMXfutYebdhoCzqEQAvD_BwE).

<sup>108</sup> См. [https://www.un.org/sites/un2.un.org/files/un\\_ai\\_advisory\\_body\\_governing\\_ai\\_for\\_humanity\\_interim\\_report.pdf](https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf). См. также <https://oecd.ai/en/accountability>.

компаниями, разработчиками ИИ и службами безопасности<sup>109</sup>. В настоящее время в мире отсутствует справедливое распределение выгод и рисков, связанных с разработкой и использованием ИИ на основе солидарности, и, как представляется, имеется существенное недоверие между организациями гражданского общества из стран глобального Юга и технологическими компаниями и разработчиками ИИ глобального Севера. Кроме того, большинство этических принципов ИИ разработаны в странах глобального Севера, и в этих принципах, возможно, не учитываются специфические проблемы глобального Юга, несмотря на то что именно глобальный Юг предоставляет полезные ископаемые и энергоресурсы, необходимые для развития инфраструктуры данных ИИ<sup>110</sup>. Наблюдается концентрация власти в руках технологических компаний и разработчиков ИИ, из-за чего повышается риск усиления мер, препятствующих проявлению солидарности в вопросах ИИ, что чревато увеличением цифрового разрыва между странами и внутри стран, а также между различными секторами общества. Форум по вопросам управления Интернетом отмечает, что существует проблема отстранения представителей глобального Юга от участия в глобальных консультациях, связанная, в частности, с недостаточным финансированием, нестабильной связью и с тем, что при выборе языка диалога предпочтение отдается английскому языку<sup>111</sup>. В будущем интеграция принципа международной солидарности на начальных этапах разработки ИИ может стать одной из ключевых ценностей для развития ИИ и внедрения механизмов надзора для обеспечения процессуальной справедливости и инклюзивного участия. В целях содействия изменению политики в направлении внедрения устойчивого ИИ, основанного на принципе международной солидарности, Независимый эксперт предлагает рекомендации, которые необходимо срочно принять в течение следующих пяти лет.

## VII. Рекомендации

57. Независимый эксперт рекомендует государствам, компаниям и субъектам гражданского общества:

а) добиваться того, чтобы государственные учреждения, технологические компании и группы гражданского общества, внедряющие ИИ для обработки данных и принятия решений, обеспечивали активное участие всех людей и групп, независимо от их расы, цвета кожи, происхождения, гендерной принадлежности, возраста, языка, религии, политических взглядов, национального или этнического происхождения, социального или экономического положения, инвалидности или любых других факторов в соответствии с Рекомендацией об этических аспектах искусственного интеллекта ЮНЕСКО;

б) принимать национальные нормативно-правовые акты, способствующие учету принципа международной солидарности на самых ранних этапах разработки ИИ, используя недискриминационный подход и обеспечивая участие гражданского общества и других соответствующих заинтересованных сторон в планировании, выборе, разработке и внедрении ИИ в государственных учреждениях и технологических компаниях, которые участвуют в обработке данных и принятии решений;

<sup>109</sup> См. <https://www.techradar.com/pro/top-tech-companies-ai-developers-and-security-firms-sign-anti-deepfake-agreement-to-combat-election-interference>.

<sup>110</sup> Материалы, представленные Форумом по вопросам управления Интернетом.

<sup>111</sup> Там же.

с) правительствам и компаниям следует постоянно проводить оценку воздействия на права человека при непосредственном участии гражданского общества до сбора, обработки или размещения данных и обеспечивать полную транспарентность, в том числе в вопросах обучения ИИ;

д) добиваться того, чтобы государственные учреждения и технологические компании, внедряющие ИИ для обработки данных и принятия решений, систематически проводили независимые, внешние аудиторские проверки на предмет транспарентности, а также оценки воздействия на права человека (для выявления элементов предвзятости в обучении, разработке алгоритмов и принятии решений), рассматривая обеспечение международной солидарности как ценность и цель на протяжении всего жизненного цикла ИИ<sup>112</sup>;

е) предусматривать в национальном и региональном законодательстве независимые, транспарентные, доступные, эффективные механизмы рассмотрения жалоб и подачи апелляций, чтобы гарантировать привлечение к ответственности за дискриминацию и/или изоляцию посредством ИИ или с помощью ИИ, который используется в государственных учреждениях, технологических компаниях или группах гражданского общества, участвующих в обработке данных или принятии решений;

ф) признать, что в случаях, связанных с дискриминацией или другими нарушениями прав человека в результате обработки данных или принятия решений с применением ИИ, бремя доказывания того, как были использованы технологии ИИ, должно лежать на соответствующих государственных учреждениях и/или технологических компаниях;

г) принять четкие правила в отношении ИИ в рамках институциональных механизмов, занимающихся оценкой воздействия обработки данных и принятия решений, а также защитой прав человека, когда разработка сопряжена с использованием персональных данных (например, данные о здоровье) или огромных объемов данных (например, национальные базы данных);

h) обеспечивать, чтобы государственные учреждения и технологические компании или группы гражданского общества предоставляли всем лицам и группам, находящимся в уязвимом положении, информацию в понятной форме о том, что при обработке данных или принятии решений будет использоваться ИИ и что запрос на получение согласия будет направляться заранее, без прямого или косвенного принуждения;

i) предоставлять доступ к независимым механизмам правосудия для решения проблемы связанных с использованием ИИ дискриминации и других нарушений прав человека;

j) правительствам и технологическим компаниям следует предоставлять группам населения, находящимся в уязвимом положении, понятную юридическую информацию об их правах и доступных им механизмах защиты в случае нарушений, связанных с использованием ИИ;

k) инвестировать в просвещение широкой общественности, направленное на повышение информационной грамотности, чтобы бороться с

<sup>112</sup> См. <https://digitalrights-check.bmz-digital.global/>;  
см. также <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>  
и <https://ieeexplore.ieee.org/document/8058187>.

дипфейками, заведомо ложной информацией и языком ненависти, сгенерированными с помощью ИИ, создавая более устойчивое общество;

l) признавать право всех людей на самореализацию в качестве одного из ключевых критериев подхода, предусматривающего учет принципа международной солидарности на самых ранних этапах разработки ИИ, а также процессы, направленные, в частности, на обеспечение доступа к образованию, жилью, рабочим местам, медицинскому обслуживанию;

m) соблюдать автономию владельцев данных при сборе и обработке данных;

n) правительствам и технологическим компаниям следует информировать владельцев данных об их правах на защиту данных и прилагать все необходимые усилия для того, чтобы у каждого человека был доступ к правам на защиту данных, включая право на их удаление, право на доступ к ним и право на отказ от предоставления данных;

o) правительствам и технологическим компаниям, обрабатывающим данные для дополнительных целей, не обозначенных изначально (вторичное использование данных), следует получать от людей свободное, предварительное и осознанное согласие, независимо от того, как эти данные были получены; органам безопасности не следует преследовать активистов, выступающих за международную солидарность, используя инструменты ИИ для борьбы с экстремизмом или терроризмом, как это предусмотрено правовыми нормами, касающимися принципов необходимости, соразмерности, законности и недискриминации;

p) разрабатывать правовые основы и систематические процедуры для определения сферы общественных интересов и классификации срочности и необходимости обработки данных; за людьми должно оставаться право отказываться от предоставления данных;

q) государствам и компаниям следует запретить создание без согласия синтетических или сфальсифицированных дипфейков, содержащих сцены интимного характера, в соответствии с концепцией разработки ИИ на основе принципа международной солидарности в отношении гендерных вопросов.

58. Независимый эксперт рекомендует гражданскому обществу:

a) продолжать выступать за инклюзивность в процессах создания этичных инструментов модерации контента с помощью ИИ для предотвращения, выявления и устранения насилия и дискриминации в Интернете в отношении женщин и лиц, подвергающихся межсекторальной дискриминации;

b) продолжать предоставлять Организации Объединенных Наций, национальным правительствам и технологическим компаниям или разработчикам ИИ примеры алгоритмической дискриминации с целью выработки рекомендаций по повышению эффективности практических методов;

c) профсоюзам трудящихся следует способствовать вовлечению работников в процессы совместной разработки, обучения и оценки рисков ИИ и проведению с ними соответствующих консультаций, а также обеспечению доступа к инструментам цифровой грамотности.

59. Независимый эксперт рекомендует компаниям:

a) создавать конкретные механизмы для распространения точек зрения традиционно маргинализированных групп, чтобы способствовать

формированию безопасной, инклюзивной онлайн-среды, основополагающим принципом которой является взаимное уважение разнообразия, в том числе проводить консультации на разных языках и предусматривать субсидии для содействия участию заинтересованных сторон из стран глобального Юга;

b) технологическим компаниям и разработчикам ИИ следует прилагать усилия, чтобы принять подход к созданию ИИ на основе принципа международной солидарности, направленный на соблюдение пределов возможностей нашей планеты по умолчанию; постоянно проводить независимые, транспарентные оценки экологических рисков; продвигать энергоэффективные алгоритмы; способствовать устойчивому цифровому развитию, внедряя практику использования возобновляемых источников энергии; обеспечивать финансирование экономики, основанной на концепции разработки ИИ с учетом принципа солидарности, для формирования данных в поддержку получения знаний об окружающей среде и изменении климата в общих зонах, доступных для общественности (таких как морское дно и океан); управлять электронными отходами; и придерживаться принципов циркулярной экономики, учитывающих интересы местных сообществ;

c) субъектам, финансирующим деятельность в области ИИ, и разработчикам ИИ следует быть приверженными концепции создания ИИ с учетом принципа международной солидарности, систематически проводя превентивные оценки воздействия на права человека на протяжении всего жизненного цикла данных для выявления риска использования алгоритмической предвзятости в нарушение принципов недискриминации и равенства, а также требовать включения прямых и косвенных заинтересованных сторон из числа гражданского общества (включая различные сообщества людей, чей голос подавляется) в независимый надзорный механизм для выявления программ, противоречащих идеям гендерного равенства и демократии, и реагирования на них;

d) технологическим компаниям следует создавать превентивные и реактивные механизмы солидарности для борьбы со связанными с использованием ИИ нарушениями прав человека, такими как кампании по распространению заведомо ложной и недостоверной информации, приводящими к насилию в обществе или к преследованию различных сообществ, чей голос подавляется, слежке за ними, их дискриминации или несоразмерно жесткой цензуре<sup>113</sup>;

e) технологическим компаниям следует создавать группы быстрого реагирования, которые обладают достаточными ресурсами и полномочиями принимать решения, чтобы своевременно реагировать на жалобы и предлагать способы урегулирования ситуации.

**60. Независимый эксперт рекомендует Организации Объединенных Наций:**

a) создать цифровую платформу солидарности, на которой организации гражданского общества (из стран глобального Севера и глобального

<sup>113</sup> Эта рекомендация соотносится с рекомендацией, содержащейся в документе «Глобальные принципы добросовестности в информационной среде Организации Объединенных Наций»: «Повышать эффективность мер реагирования на кризисные ситуации. Сотрудничая с заинтересованными сторонами, действующими в зонах повышенного риска, разрабатывать процедуры раннего предупреждения и передачи по инстанции с ускоренным и своевременным реагированием в условиях кризиса и конфликта. Создавать механизмы, обеспечивающие свободный и своевременный доступ к надежной и точной информации, отвечающей общественным интересам».



**Юга) могли бы обмениваться идеями относительно международной солидарности, а также обсуждать с представителями технологических компаний, разработчиками ИИ и государственными чиновниками стратегии и практические методы, проблемы и инновационные решения, касающиеся соблюдения принципов инклюзивности и недискриминации в области ИИ, в соответствии с рекомендациями, изложенными в докладе Независимого эксперта о гражданском обществе и международной солидарности<sup>114</sup>;**

**б) учредить фонд в поддержку создания больших языковых моделей для данных на всех языках в целях сохранения культурного разнообразия.**

---

<sup>114</sup> A/HRC/56/57.