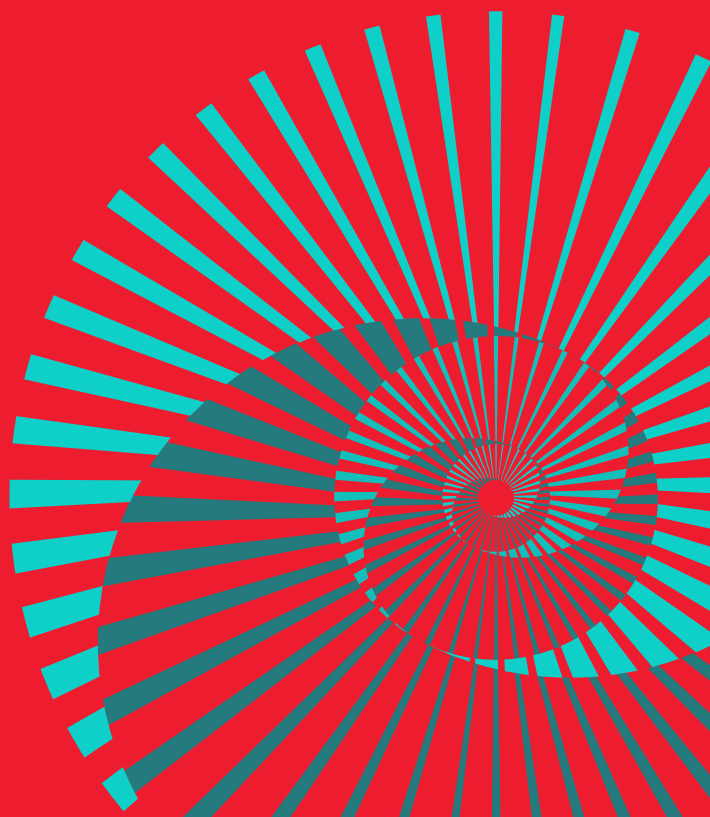


Размышления на тему: свобода мысли в эпоху искусственного интеллекта



АНАЛИТИЧЕСКАЯ
ЗАПИСКА

ОБСЕ
БДИПЧ



Размышления на тему: свобода мысли в эпоху искусственного интеллекта

АНАЛИТИЧЕСКАЯ ЗАПИСКА



Благодарность

БДИПЧ ОБСЕ выражает искреннюю благодарность всем, кто щедро поделился своими знаниями и опытом в ходе подготовки этой публикации. Их содержательные рецензии и отзывы на разных этапах работы внесли неоценимый вклад в этот процесс. БДИПЧ выражает особую благодарность членам Группы экспертов БДИПЧ по свободе религии и убеждений, в особенности Ахмеду Шахиду, Дэвиду Гриффитсу и Паскуале Анниккино за их обстоятельные комментарии. БДИПЧ также выражает признательность Камерану Ашрафу, Кристофу Бублицу, Эрве Шнейвайсу, Эмили Эльштуб, Александру Крибицу и Саймону Маккарти-Джонсу, а также нашим коллегам из аппарата Представителя ОБСЕ по вопросам свободы СМИ за рецензирование различных вариантов проекта данной публикации и ценные предложения. БДИПЧ также выражает особую благодарность Бенджамину Гринакру, чьи работы по данной теме в значительной степени помогли улучшить окончательный вариант документа, и Милене Костас Траскасас за постоянный диалог и дух сотрудничества.

Размышления на тему: свобода мысли в эпоху искусственного интеллекта

Опубликовано Бюро ОБСЕ по демократическим институтам и правам человека (БДИПЧ)

ul. Miodowa 10

00-251 Warsaw

Poland

www.osce.org/odihr

© БДИПЧ ОБСЕ, 2026 Все права защищены.

Содержание данной публикации можно свободно использовать и копировать в образовательных и других некоммерческих целях при условии, что любое такое воспроизведение будет сопровождаться указанием БДИПЧ ОБСЕ в качестве источника.

ISBN: 978-92-9271-490-1

Оформление: Филип Андроник (Filip Andronik)



Оглавление

Резюме	5
Введение	7
1. О свободе мысли	9
2. Об искусственном интеллекте	16
Генеративный ИИ.....	17
Влияние корпораций в эпоху ИИ	20
3. Алгоритмическое мышление и его последствия для прав человека	23
ИИ, предвзятость и мышление.....	23
Человеческий контроль, прозрачность и подотчетность	27
4. Свобода мысли в изменившемся гражданском пространстве	30
Как новая архитектура гражданского пространства влияет на свободу мысли.....	30
Как «информационная путаница» меняет реальный выбор	36
Как ИИ влияет на связи между людьми, развитие мышления и право на забвение	39
5. Нейротехнологии и свобода мысли	43
6. Нормативно-правовая база	50
7. Выводы и рекомендации	58

Резюме

Настоящая аналитическая записка представляет собой краткое введение в тему права на свободу мысли, в котором отмечается, что на государствах лежит позитивная обязанность по созданию благоприятных условий для его полного осуществления¹. В начале свобода мысли рассматривается как основополагающее право, которое необходимо защищать и поддерживать в любых контекстах. Далее говорится о том, как новые технологии – искусственный интеллект (ИИ) и нейротехнологии – меняют условия реализации свободы мысли. При том что эти системы и технологии (в зависимости от того, как они разрабатываются и внедряются) могут использоваться для расширения возможностей для реализации прав человека², в частности, путем улучшения доступа к разнообразной информации, оказания содействия образовательным процессам или развитию медицины, они вместе с тем являются источником беспрецедентных рисков. Перед разработкой, внедрением и использованием таких технологий необходимо проводить комплексную оценку воздействия и предусматривать надежные защитные механизмы.

В данной аналитической записке рассматриваются последствия новой информационной экосистемы для свободы мысли, а также потенциальное воздействие нейротехнологий на основе ИИ, в связи с которыми возникают фундаментальные вопросы о неприкосновенности «ментальной автономии» личности. В аналитической записке также излагаются рекомендации для государств по обеспечению свободы мысли в цифровую эпоху, в том числе по следующим аспектам:

- регуляторные и институциональные защитные механизмы для противодействия возникающим рискам, особенно связанным с искусственным интеллектом и нейротехнологиями;

1 Хотя в документе признается, что некоторые направления использования ИИ в военной области и в сфере наблюдения имеют существенные последствия для свободы мысли, эти вопросы не рассматриваются в данной аналитической записке.

2 Theresa Addie, MS., [Harnessing Technology to Safeguard Human Rights: AI, Big Data, and Accountability](#) [Технологии на страже прав человека: ИИ, большие массивы данных и подотчетность], Специальный докладчик ООН по вопросам свободы религии и убеждений Ахмед Шахид, [Доклад Специального докладчика по вопросам свободы религии и убеждений](#), A/76/380, 5 октября 2021 г., пункты 2-4.

- механизмы этического надзора и обязательная оценка воздействия на права человека, предусмотренные на всех этапах жизненного цикла ИИ³;
- инициативы по обеспечению прозрачности, инклюзивных общественных обсуждений и подотчетности в рамках процессов принятия решений;
- развитие потенциала разработчиков и специалистов по системам ИИ посредством подготовки и обучения по вопросам этики и прав человека;
- усиление роли независимых институтов в области защиты свободы мысли и демократических ценностей.

Авторы аналитической записки приходят к выводу, что инновации, изначально не учитывающие права человека и не обеспечивающие их защиту, подрывают цели, призванные служить ориентиром для этического, научного и технологического прогресса, а именно — обеспечение пользы для человечества⁴.

Исходя из работы, проделанной международными организациями и научными кругами в области ИИ, нейротехнологий и прав человека в данной публикации рассматриваются соответствующие конкретные последствия для свободы мысли. Этот документ призван стать основой для дальнейшего обсуждения и источником информации для инициатив и действий БДИПЧ, государств-участников и других заинтересованных сторон.

3 [Глобальный цифровой договор ООН](#), резолюция Генеральной Ассамблеи ООН A/RES/79/1 «Пакт во имя будущего» от 22 сентября 2024 года, охватывает следующие этапы жизненного цикла цифровых и новых технологий: предварительное проектирование, проектирование, разработка, оценка, тестирование, внедрение, использование, продажи, закупки, эксплуатация и вывод из эксплуатации.

4 [Декларация об использовании научно-технического прогресса в интересах мира и на благо человечества](#), принята Резолюцией 3384 (XXX) Генеральной Ассамблеи от 10 ноября 1975 года.

Введение

По мере того, как правительства осваивают стремительно развивающуюся сферу искусственного интеллекта (ИИ), в частности, его интеграцию в такие области, как нейротехнологии, важнейшее, но часто игнорируемое, право человека — право на свободу мысли — требует безотлагательного внимания. Это право закреплено во Всеобщей декларации прав человека (ВДПЧ) наряду с правом на свободу совести и вероисповедания и гарантирует способность человека мыслить самостоятельно, формировать свои собственные убеждения, принимать независимые решения и предлагать новые идеи. Это право поддерживает креативность, инновации и формирование устойчивых и плюралистических обществ⁵.

Кроме того, это право предусматривает свободу не раскрывать свои мысли, а также свободу от принуждения, наказания за мысли и недопустимого изменения мыслей человека⁶. Признавая основополагающий характер этого права, государства-участники ОБСЕ взяли на себя обязательство защищать свободу мысли. Их первым шагом стало принятие Хельсинкского Заключительного акта⁷.

5 [Международный пакт о гражданских и политических правах](#), принят резолюцией 2200 А (XXI) Генеральной Ассамблеи от 16 декабря 1966 года, ст. 18:

- «1. Каждый человек имеет право на свободу мысли, совести и религии. Это право включает свободу иметь или принимать религию или убеждения по своему выбору и свободу исповедовать свою религию и убеждения как единолично, так и сообща с другими, публичным или частным порядком, в отправлении культа, выполнении религиозных и ритуальных обрядов и учении.
2. Никто не должен подвергаться принуждению, умаляющему его свободу иметь или принимать религию или убеждения по своему выбору.
3. Свобода исповедовать религию или убеждения подлежит лишь ограничениям, установленным законом и необходимым для охраны общественной безопасности, порядка, здоровья и морали, равно как и основных прав и свобод других лиц.
4. Участвующие в настоящем Пакте Государства обязуются уважать свободу родителей и в соответствующих случаях законных опекунов, обеспечивать религиозное и нравственное воспитание своих детей в соответствии со своими собственными убеждениями».

6 Это относится к любому изменению мыслей способом, противоречащим стандартам прав человека, в том числе посредством воздействия на мозг. См. Ахмед Шахид, [Доклад Специального докладчика по вопросам свободы религии и убеждений](#).

7 [Заключительный акт](#) первого Совещания по безопасности и сотрудничеству в Европе (СБСЕ), 1 августа 1975 г, признал одним из основных руководящих принципов «уважение прав человека и основных свободы, включая свободу мысли, совести, религии и убеждений». Это право было закреплено в ряде документов ОБСЕ (например, [Итоговый документ Венской встречи 1986 года](#),

Вместе с тем на фоне стремительного развития новых технологий, включая искусственный интеллект и нейротехнологии, реализация права на свободу мысли оказалась под угрозой. Поскольку такие системы и технологии все больше влияют на то, как мы думаем, общаемся и взаимодействуем, возникают сложные юридические, этические и социальные вопросы, которые требуют от государств-участников внимания и действий.

4 ноября 1986 г. по 19 января 1989 г.; [Документ Копенгагенского совещания конференции по человеческому измерению СБСЕ](#), 29 июня 1990 г.; [Решение № 4/03 Совета министров ОБСЕ «Толерантность и недискриминация»](#) [Маастрихт, 2 декабря 2003 г. (на англ. языке)]; [Решение № 3/13 Совета министров ОБСЕ «Свобода мысли, совести, религии и убеждений»](#), Киев, 9 декабря 2013 г.

1. О свободе мысли

Статья 18 Всеобщей декларации прав человека (ВДПЧ) защищает право на «свободу мысли, совести и религии»⁸. Как правило, основной акцент делается на религии: Статью 18 часто называют сокращенно статьей «о свободе религии или убеждений» или «о свободе вероисповедания». До недавнего времени неотъемлемая защита свободы мысли считалась настолько очевидной, что ее определение казалось ненужным, несмотря на то что она является основополагающим началом человеческого достоинства и автономии, а также всех других прав человека. Один из авторов Всеобщей декларации прав человека Рене Кассен называл свободу мысли «источником всех других прав», но ни он, ни другие авторы декларации не излагали свое видение более подробно, при этом судебная практика, касающаяся свободы мысли, является незначительной⁹.

В то же время после Второй мировой войны авторы Всеобщей декларации прав человека хорошо понимали, что может угрожать этому праву. Попытки повлиять на мышление или изменить его, а также наказывать за реальные или предполагаемые мысли и мнения имеют давнюю историю¹⁰. Люди, взгляды которых расходятся с общепринятыми, которые мыслят иначе и подвергают сомнению действующие нормы, сталкиваются с принудительными мерами, которые, как правило, ужесточаются в условиях автократии¹¹.

⁸ [Всеобщая декларация прав человека](#), Резолюция ГА ООН 217 А, 10 декабря 1948 г.

⁹ Ахмед Шахид, [Доклад Специального докладчика](#).

¹⁰ От нацистской пропаганды, которая привела к Холокосту, до тоталитаризма и авторитаризма, а также других ситуаций, когда государства предпринимали действия, препятствующие свободе мысли, например, меры по предотвращению насильственного экстремизма и радикализации, ведущих к терроризму. См., например, [‘This is the Thought Police’: The Prevent Duty and Its Chilling Effect on Human Rights](#) [Это полиция мысли: обязанность по предотвращению и ее негативное воздействие на права человека], (London: Amnesty International, 2023); Специальный докладчик ООН по вопросу о свободе религии или убеждений Ахмед Шахид, [Доклад о взаимосвязи между свободой религии или убеждений и насильственным экстремизмом](#), A/73/362, 5 сентября 2018 г., пункт 43.

¹¹ По состоянию на 2023 год 71% населения мира проживал в странах с авторитарным режимом, в 2003 году эта доля составляла 50%; в 2024 году 42 страны (в которых проживает 35% населения мира) переживали процесс «автократизации», а 18 стран (в которых проживает 5% населения мира) — процесс демократизации. См. Marina Nord, Martin Lundstedt, and Staffan I Lindberg, [Media Freedom, Democracy, and Security, Research Report](#) [Доклад о свободе СМИ, демократии и безопасности], OSCE/RFoM, 15 July 2024, p. 6

Несмотря на отсутствие в международном праве определения понятия «мысль»¹², можно определить его контуры и признаки, чтобы обеспечить его защиту на практике.

В соответствии со Статьей 18 Всеобщей декларации прав человека и Международным пактом о гражданских и политических правах (МПГПП) это право обычно рассматривается в контексте свободы совести и свободы религии или убеждений. Вместе с тем свобода мысли не ограничивается вопросами религиозной свободы. Комитет ООН по правам человека дает авторитетное толкование Статьи 18 МПГПП в Замечании общего порядка № 22, где поясняет, что она «охватывает свободу мысли по всем вопросам»¹³.

Следует отметить, что отступление от Статьи 18 не допускается: это означает, что государства не могут приостанавливать выполнение своих обязательств в отношении этого права даже в условиях «чрезвычайного положения в государстве, при котором жизнь нации находится под угрозой»¹⁴. Свобода мысли обычно понимается как абсолютное и безусловное право. В Замечании общего порядка № 22 отмечается, что Статья 18 «не допускает установления каких-либо ограничений свободы мысли и совести или свободы иметь или принимать религию или убеждения по своему выбору»¹⁵. Эти свободы, наряду с правом беспрепятственно придерживаться своих мнений (Статья 19.1 МПГПП), образуют то, что в праве в области прав человека называют *forum internum*, или внутреннее пространство человека (разум)¹⁶.

Forum internum противопоставляется *forum externum*, который касается выражения мыслей, убеждений или веры. Права *forum externum* не являются абсолютными¹⁷. Напротив, они могут быть ограничены

12 Ахмед Шахид, Доклад Специального докладчика, п. 11.

13 Замечание общего порядка № 22, Статья 18 (Свобода мысли, совести и религии), Комитет ООН по правам человека, 48-я сессия, 27 сентября 1993 г., п. 1.

14 Не подлежит отмене и ряд других прав, в том числе Статья 6 (право на жизнь); Статья 7 (право не подвергаться пыткам или жестокому, негуманному или унижающему достоинство обращению или наказанию); Статья 8, пункты 1 и 2 (право не находиться в рабстве или подневольном состоянии и обязанность запрещать рабство и работорговлю во всех их формах).

15 Ахмед Шахид, Доклад Специального докладчика, п. 3.

16 Там же, п. 2.

17 МПГПП, ст. 18.3: Свобода исповедовать религию или убеждения подлежит лишь ограничениям, установленным законом и необходимым для охраны общественной безопасности, порядка, здоровья и морали, равно как и основных прав и свобод других лиц. МПГПП, ст. 19.3: Пользование

государством, «как установлено законом и необходимо для охраны государственной безопасности, общественного порядка, здоровья или нравственности населения, равно как и основных прав и свобод других лиц» (а в случае Статьи 19.3 — по дополнительным основаниям защиты национальной безопасности).

В этой формулировке проводится различие между аспектами Статей 18 и 19, которые имеют абсолютную защиту, и теми, которые могут быть ограничены при определенных обстоятельствах. Вместе с тем рассмотрение свободы мысли исключительно в контексте *forum internum* наделяет ее пассивным характером. Мысль — это не только движущая сила деятельности человека, но и сама деятельность, которая должна осуществляться беспрепятственно.

Как и все другие права человека, свобода мысли часто пересекается с другими правами и свободами на фоне их глубокой взаимосвязи и взаимозависимости. Пожалуй, одним из важнейших является ее связь с неприкосновенностью частной жизни, правом, которое в отдельных ситуациях может быть ограничено, с точки зрения защиты (ментального) пространства для рождения мысли. Такое пересечение также иллюстрирует необходимость дальнейшего определения рамок свободы мысли и более четкого разъяснения ее защиты, особенно с учетом абсолютного характера этого права. Например, при каких обстоятельствах неприкосновенность частной жизни и неприкосновенность мыслительной деятельности предусматривают абсолютную защиту в рамках права на свободу мысли? И когда достаточно было бы квалифицированной защиты в рамках права на неприкосновенность частной жизни¹⁸?

предусмотренными в пункте 2 настоящей статьи правами налагает особые обязанности и особую ответственность. Оно может быть, следовательно, сопряжено с некоторыми ограничениями, которые, однако, должны быть установлены законом и являться необходимыми:

- a) для уважения прав и репутации других лиц;
- b) для охраны государственной безопасности, общественного порядка, здоровья или нравственности населения.

18 Научная литература предлагает ряд критериев, которые помогают различать ситуации, требующие абсолютной и ограниченной защиты в отношении свободы мысли и неприкосновенности частной жизни соответственно. Один из критериев учитывает ментальные последствия вмешательства (главным образом с точки зрения их «тяжести» и «значимости») наряду с другими релевантными факторами, такими как значение личной заинтересованности в неразглашении определенных типов ментального содержания (например, сексуальная ориентация, политические взгляды, религиозные убеждения) или метод вмешательства, посредством которого раскрывается или изменяется ментальное состояние человека (например, индоктринация или манипулирование).

Кроме того, свободу мысли следует рассматривать в контексте свободы совести и свободы религии или убеждений. Совесть, как «внутренний суд»¹⁹, позволяет нам на самом базовом уровне свободно решать, что является правильным, а что неправильным. При этом так происходит только в том случае, если она основана на независимом мышлении, то есть автономия каждого человека в формировании своих мыслей не подвергается неоправданному вмешательству. Точно так же имеется риск того, что приверженность любой религии или системе взглядов может по своей природе быть принудительной в отсутствие свободы мысли.

Процесс зарождения мысли происходит не в вакууме, а во взаимодействии с другими людьми, пространствами и информацией, где важную роль играют сообщества. Это значит, что существует значимое пересечение с правами, закрепленными в Статье 19 «О свободе мнений и их свободном выражении»²⁰. Мысль и ее выражение в своем взаимодействии «образуют неразрывный замкнутый круг»²¹. При этом они не идентичны и пользуются разной защитой в рамках международного права в области прав человека. Мысль — от стимула до результата — относится к понятию мыслительной

Уровень контроля над собственными мыслями после вмешательства также имеет значение, как и то, подрывает ли этот метод такой контроль или позволяет обойти его. Другой критерий рассматривает характеристики пострадавшего с точки зрения уязвимости; особенно уязвимы дети, пожилые люди, люди с психическими расстройствами, лица, лишённые свободы или находящиеся в иных условиях институционального контроля. В этом смысле важны различия в уровне власти, в том числе в ситуациях на рабочем месте. Ещё одним критерием является контекст, в котором произошло вмешательство, что требует оценки в каждом конкретном случае. См. Sjors Lighthart and Naomi van de Pol, *Freedom of Thought: Absolute Protection of Mental Privacy and Mental Integrity? Considering the Case of Neurotechnology in Criminal Justice* [Свобода мысли: абсолютная защита ментальной неприкосновенности и конфиденциальности ментальной информации? Нейротехнологии в уголовном правосудии], in Patrick O'Callaghan and Bethany Shiner (eds.), *The Cambridge Handbook of the Right to Freedom of Thought*, (Cambridge: Cambridge University Press, 2025).

19 Sofie Møller, *Kant's Tribunal of Reason, Legal Metaphor and Normativity in the Critique of Pure Reason* [Кантовский суд разума, правовая метафора и нормативность в «Критике чистого разума»], (Cambridge: Cambridge University Press, 2020), Chapter 6 - **Moral Conscience as the Practical Inner Tribunal** [Совесть как действенный внутренний суд] pp. 96-112.

20 К другим релевантным правам относятся следующие: право каждого человека на пользование наивысшим достижимым уровнем физического и психического здоровья (Статья 12 Международного пакта об экономических, социальных и культурных правах (МПЭСКП)); право каждого человека на образование — в том числе в том смысле, что «образование должно быть направлено на всестороннее развитие личности и чувства собственного достоинства» (Статья 13 МПЭСКП); право участвовать в культурной жизни и пользоваться плодами научного прогресса и его применения; право на защиту моральных и материальных интересов, вытекающих из любого научного, литературного или художественного произведения, автором которого он является, или государства, обязующиеся уважать свободу, необходимую для научных исследований и творческой деятельности (Статья 15 МПЭСКП); или запрет принуждать кого-либо к членству в ассоциации (Статья 20.2 ВДПЧ).

21 Ахмед Шахид, *Доклад Специального докладчика*, п. 18.

деятельности, тогда как мнение можно рассматривать как один из ее предварительных результатов²².

«<...> человек развивается и добивается успеха в процессе взаимодействия с другими людьми и в благоприятных с материальной и культурной точек зрения условиях, <...> независимость носит не только индивидуальный характер, она зависит от отношений, поскольку возникает в результате взаимодействия с сообществом и на основе чувства принадлежности к нему».

— Проект рекомендации об этических аспектах нейротехнологий, ЮНЕСКО, 9 апреля 2025 г., пп. 16 и 44.

В основополагающем докладе 2021 года (первом в рамках системы ООН содержательном комментарии о праве на свободу мысли) Специальный докладчик ООН по вопросу о свободе религии или убеждений Ахмед Шахид рассматривает четыре предлагаемых атрибута этого права:

- свобода не раскрывать свои мысли;
- свобода от наказания за свои мысли, реальные или предполагаемые;
- защита от недопустимого изменения другими своих мыслей;
- позитивное обязательство государств создавать условия, благоприятствующие осуществлению свободы мысли.

В то время как первые два пункта являются более доступными для понимания, защита от недопустимого изменения мысли обычно подразумевает запрет неправомерного вмешательства в «ментальную автономию», т. е. отсутствие принуждения и изменения чье-либо мышления посредством воздействия на мозг и манипулирования.

²² Там же, п. 21; см. также Christoph Bublitz, 'The Mind and Conscience are the Person's Most Sacred Possessions': The Origins of Freedom of Thought in the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights [Разум и совесть — священные достояния человека. Истоки свободы мысли во Всеобщей декларации прав человека и Международном пакте о гражданских и политических правах], in Patrick O'Callaghan and Bethany Shiner (eds.), *The Cambridge Handbook of the Right to Freedom of Thought*, (Cambridge: Cambridge University Press, 2025

В то же время свобода мысли не гарантирует защиту от влияния мыслей других людей или от процессов убеждения в повседневной жизни. Специальный докладчик также предлагает среди прочих факторов критерии для выявления незаконных форм манипуляции, которые необходимо оценивать в каждом конкретном случае:

- согласие (добровольное и осознанное, полученное от лица, способного его дать);
- сокрытие и запутывание (если «разумный человек» в состоянии понять, что на него оказывается влияние);
- асимметричность властных полномочий между субъектом, оказывающим влияние, и носителем права и то, как она использовалась;
- вред как цель или результат.

Четвертый атрибут — обязанность создавать условия, благоприятствующие осуществлению свободы мысли — связан с другими правами, включая право на доступ к информации, коммуникации и образованию²³.

«Нам нужна свобода мысли, чтобы бороться с изменением климата, расизмом и глобальной бедностью, а также чтобы влюбляться, смеяться и мечтать. Право на свободу мысли является индивидуальным правом, но оно имеет решающее значение для культурной, научной, политической и эмоциональной жизни наших обществ. Свобода мысли дает нам возможность думать о неприятных вещах и отгонять эти мысли, прежде чем мы начнем действовать в соответствии с ними или позволим им укорениться; она позволяет нам выбирать, как мы будем вести себя по отношению к другим, адаптировать нашу речь в зависимости от контекста и аудитории и быть самими собой. Свобода мысли позволяет нам представлять себе новое будущее, не доказывая его сначала, она сохраняет нашу динамичность и склонность к приключениям, она обеспечивает

23 Ахмед Шахид, Доклад Специального докладчика, пп. 25–47.

нам безопасность и прежде всего позволяет нам оставаться людьми».

— Susie Alegre, *Freedom to Think: Protecting a Fundamental Human Right in the Digital Age* [Свобода мысли: защита одного из основных прав человека в цифровую эпоху], (London: Atlantic Books, 2022), Introduction.

2. Об искусственном интеллекте

В международном праве нет общепринятого определения ИИ. В [Законе Европейского союза \(ЕС\)](#) об искусственном интеллекте²⁴ «система ИИ» определена как «машинная система, предназначенная для работы с различными уровнями автономии, которая может проявлять адаптивность после развертывания и которая для достижения явных или неявных делает выводы на основе входных данных о том, как формировать выходные данные, включая прогнозы, контент, рекомендации и решения, которые могут воздействовать на физическую или виртуальную среду»²⁵. Аналогичное определение содержится в Конвенции Совета Европы об искусственном интеллекте и правах человека, демократии и верховенстве права²⁶.

В основе ИИ лежат алгоритмы, реализованные в программном коде²⁷, которые обрабатывают наборы данных путем сбора и преобразования исходных данных для формирования выходного результата²⁸. В отличие от алгоритма, где логика принятия решений заранее задана, ИИ «принимает решения» на основе анализа данных²⁹. Он выявляет закономерности в данных, «обучается» на них и делает прогнозы, которые затем может адаптировать по мере получения новых данных³⁰.

Потенциальные сферы применения ИИ практически не ограничены, и его последствия для жизни людей могут быть самыми разными.

24 [Закон Европейского союза \(ЕС\)](#) 2024/1689 об искусственном интеллекте, официальный журнал, в редакции от 13 июня 2024 г. (на англ. языке).

25 Там же, Ст. 3 (1).

26 [Рамочная конвенция об искусственном интеллекте и правах человека, демократии и верховенстве права](#), Совет Европы, принята 5 сентября 2024 г. (на англ. языке).

27 Алгоритмы — это набор инструкций, направленных на решение конкретной проблемы. Процессы, на базе которых они работают, являются сложными и разнообразными. См., например, Andrew Williams, [What Is an Algorithm? Defining And Applying Algorithms](#) [Что такое алгоритм? Определение и применение алгоритмов], *Forbes*, 12 January 2024

28 Специальный докладчик ООН по вопросам поощрения и защиты права на свободу мнений и их свободное выражение Дэвид Кей, [A/73/348](#), 29 августа 2018 г.

29 Kaya Ismail, [AI vs. Algorithms: What's the Difference?](#) [Искусственный интеллект и алгоритмы: в чем разница?], *CMSWire*, 26 October 2016.

30 Office of Communications, College of Education, [Traditional AI vs. Generative AI: What's the Difference?](#) [Традиционный ИИ и генеративный ИИ: в чем различие?], University of Illinois Urbana-Campaign, 11 November 2024

Искусственный интеллект работает с наборами данных, которые отражают социальные реалии и персональные данные, в том числе в таких областях, как здравоохранение, занятость, банковская сфера, торговля и правоохранительная деятельность. Например, в здравоохранении ИИ способен анализировать наборы медицинских данных для прогнозирования результатов лечения, на основе которых принимаются медицинские решения. ИИ также используется для анализа больших наборов данных о покупательском поведении онлайн и выявления индивидуальных предпочтений. Он может сопоставить эту информацию с персональными и контекстными характеристиками (такими как возраст, географическое положение, пол, прошлые покупки и покупательское поведение), для того чтобы формировать детализированные пользовательские профили. Компании используют такие профили для рыночной сегментации, что позволяет им предлагать людям персонализированные продукты или услуги с целью увеличения продаж или роста вовлеченности. Вовлеченность позволяет собирать новые пользовательские данные.

Несмотря на то что пользователи в той или иной степени осознают, что их данные используются, а на их выбор оказывают влияние, эти процессы все равно вызывают вопросы о свободе мысли и потенциальных возможностях для манипуляций. Вместе с тем у пользователей недостаточно информации для того, чтобы участвовать в обсуждении того, как новые технологии можно использовать в общественных интересах.

Генеративный ИИ

Генеративный ИИ (GenAI) — это новое поколение разработок в области искусственного интеллекта. Это искусственный интеллект, который «создает» контент, и используется главным образом при реагировании на запросы, предоставляемые через интерфейсы на естественном языке³¹. Классические системы ИИ выдают результаты на основе имеющегося контента, тогда как GenAI генерирует новый контент

31 Fengchun Miao, Wayne Holmes, [Guidance for generative AI in education and research](#) [Руководство по генеративному ИИ в образовании и исследованиях], (Paris: UNESCO, 2023)

на основе существующего³². Это может быть текст, изображение, видео, музыка и даже программный код. Генеративный ИИ обучается на огромных наборах данных. Для генерации ответов он проводит сложный анализ закономерностей и взаимосвязей данных, таких как последовательность слов или расположение пикселей³³. Он разрабатывался и функционирует на основе так называемых больших языковых моделей и технологий глубокого обучения³⁴, которые разрабатывались и совершенствовались на протяжении нескольких лет, прежде чем ChatGPT стал общедоступным и все получили возможность экспериментировать с ним или использовать его для работы.

После того как компания OpenAI открыла доступ к ChatGPT, своей модели генеративного ИИ, обученной на данных из Интернета, эта программа опередила все аналогичные продукты по темпам роста масштабов ее распространения. Это привело к росту конкуренции среди крупных технологических компаний в сфере разработки инструментов искусственного интеллекта и их интеграции в различные продукты и платформы³⁵. Именно из-за их быстрого внедрения наблюдается стремительный рост контента, созданного с помощью ИИ, который, дополняя контент, созданный человеком, меняет характер данных в Интернете и, как следствие, доступные знания³⁶.

Один из вопросов, которые вызывают беспокойство, в том числе в отрасли, связан с тем, что, как показали исследования, при многократном обучении на искусственно сгенерированных данных в отличие от данных, сгенерированных человеком, происходит коллапс модели генеративного ИИ. Тем не менее эта ситуация может измениться с появлением новых технологических достижений³⁷.

32 Office of Communications, College of Education, [Traditional AI vs. Generative AI: What's the Difference?](#) [Традиционный ИИ и генеративный ИИ: в чем различие?]

33 Miao and Holmes, [Guidance for generative AI in education and research](#). [Руководство по генеративному ИИ в образовании и исследованиях], p. 8

34 Определение **большой языковой модели** в Кембриджском онлайн-словаре английского языка.

35 Примеры моделей GenAI: Claude (Anthropic), Llama (Meta), Gemini (Google), Le Chat (Mistral) и другие. Основные разработки проводятся крупными компаниями, при этом не только технологическими гигантами.

36 См. Jason Koebler, Project Analyzing Human Language Usage Shuts Down Because 'Generative AI Has Polluted the Data' [Проект по анализу использования человеческого языка закрыт, потому что генеративный ИИ загрязнил данные]. 404media, 19 September 2024.

37 Ilya Shumailov, Zakhar Shumaylov, et al., [AI models collapse when trained on recursively generated data](#) [Коллапс ИИ-моделей при обучении на рекурсивно сгенерированных данных]., *Nature*, 631, 24 July

«Коллапс модели — это дегенеративный процесс, влияющий на поколения обученных генеративных моделей, при котором генерируемые ими данные в конечном итоге загрязняют обучающий набор следующего поколения. При обучении на загрязненных данных они неправильно воспринимают реальность».

— Ilya Shumailov, Zakhar Shumaylov, et al., [AI models collapse when trained on recursively generated data](#) [Коллапс ИИ-моделей при обучении на рекурсивно сгенерированных данных], *Nature*, 631, 24 July 2024, p. 755-759

Китайская модель генеративного ИИ DeepSeek, которая появилась в начале 2025 года, обеспечивает высокую техническую производительность при существенно более низких издержках, в частности, с точки зрения воздействия на окружающую среду. DeepSeek имеет открытый исходный код, что позволяет загружать, копировать и использовать его без значительных затрат на создание модели с нуля³⁸. Это позволило сделать новые научные разработки доступными для всех³⁹. В то же время его появление вызвало опасения относительно прав, связанных с неприкосновенностью частной жизни и цензурой⁴⁰. Ряд стран ЕС запретили его использование в связи с рисками в области защиты данных⁴¹. Влияние DeepSeek на конкуренцию в отрасли и распространение моделей генеративного ИИ еще не проявилось в полной мере⁴².

2024, pp. 755-759.

38 Charlotte Edmond, What is open-source AI and how could DeepSeek change the industry [Что такое ИИ с открытым исходным кодом, и как DeepSeek может изменить отрасль], *World Economic Forum*, 5 February 2025.

39 Alex He, DeepSeek and China's AI Innovation in US-China Tech Competition [DeepSeek и китайские инновации в области ИИ в контексте технологической конкуренции между США и Китаем], *Centre for International Governance Innovation*, 11 April 2025.

40 Robert Booth and Dan Milmo, Experts urge caution over use of Chinese AI DeepSeek [Эксперты призывают к осторожности при использовании китайского ИИ DeepSeek], *The Guardian*, 28 January 2025.

41 Hakan Ersen and Miranda Murray, DeepSeek faces ban from Apple, Google app stores in Germany [В Германии DeepSeek может быть запрещен в магазинах приложений Apple и Google], *Reuters*, 27 June 2025.

42 Hamilton Mann, [Seek Deeper On DeepSeek For Artificial Integrity Over Intelligence](#) [DeepSeek: в поисках прозрачности, а не интеллекта], *Forbes*, 28 January 2025; Robert Booth and Dan Milmo, [Chinese AI chatbot DeepSeek censors itself in realtime, users report](#) [Китайский чат-бот с ИИ DeepSeek цензурирует себя в режиме реального времени, как сообщают пользователи], *The Guardian*, 28

Крупные технологические компании также говорят о конкретных шагах на пути к следующему этапу развития искусственного интеллекта — общему искусственному интеллекту (AGI), при этом неизвестно, насколько они действительно приблизились к его созданию⁴³. Ожидается, что AGI будет автономно решать задачи и обладать самостоятельностью на уровне человека⁴⁴. Это будет иметь последствия для свободы мысли, поскольку такие системы могут как оказывать влияние на осуществление этого права, поддерживать его или подрывать его реализацию. В этой связи необходимо, чтобы правительства, разработчики и регулирующие органы гарантировали, что создание и управление общим искусственным интеллектом будет соответствовать принципам прав человека.

Влияние корпораций в эпоху ИИ

Экономические и нормативно-правовые условия, в которых работали новые технологии в последние десятилетия, оказали значительное влияние на демократию и права человека. Использование ИИ крупнейшими технологическими корпорациями (BigTech), которым принадлежат ведущие интернет-платформы, включая поисковые системы и социальные сети, существенно изменило пространство для общественного обсуждения. Такая трансформация повлияла на демократическую дискуссию и на то, как люди формируют свое мнение, делают выбор и принимают решения, что напрямую повлияло на свободу мысли⁴⁵. На фоне того, как корпорации и правительства

January 2025; Cade Metz, [How Did DeepSeek Build Its A.I. With Less Money?](#) [Как DeepSeek создал свой ИИ дешевле], *The New York Times*, 12 February 2025.

43 Kate Brennan, Amba Kak and Dr. Sarah Myers West, *Artificial Power: 2025 Landscape Report* [Способности ИИ: обзор за 2025 год], AI Now Institute, 3 June 2025

44 Tom Allen, Real-world agentic AI is 'complex' cautions Google's Demis Hassabis [«ИИ, обладающий самостоятельностью, является сложным инструментом», — предупреждает Демис Хассабис из Google], *Computing*, 17 March 2025; Will Knight, Google's AI Boss Says Gemini's New Abilities Point the Way to AGI [«Руководитель направления по разработке ИИ в компании Google говорит, что новые возможности Gemini указывают путь к общему искусственному интеллекту»], *Wired*, 28 May 2025; Robert Booth, Meta to announce \$15bn investment in bid to achieve computerised 'superintelligence' [«Meta объявит об инвестициях в размере 15 млрд долларов в попытке создать компьютеризированный «сверхинтеллект»], *The Guardian*, 11 June 2025.

45 Подробнее см. Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* [Эпоха надзорного капитализма: борьба за будущее человечества на новом рубеже власти], (New York, Public Affairs, 2019). См. также Представитель ОБСЕ по вопросам свободы СМИ [об искусственном интеллекте и свободе выражения мнений](#).

получают все более широкий доступ к большим объемам данных о поведении и предпочтениях людей, расширяются и их возможности по оказанию влияния на мысли людей и даже манипулирования ими. Системы ИИ (за исключением лишь некоторых), независимо от того, какой используется источник финансирования (частный или государственный), сегодня не регулируются на общенациональном уровне нормативными документами, основанными на правах человека и предусматривающими предварительную оценку воздействия или их ответственность в будущем.

В то же время происходит стремительное развитие нейротехнологий, которые благодаря интеграции с ИИ становятся все более сложными. Поскольку эти технологии затрагивают работу мозга и нейронных цепей, они позволяют извлекать информацию о мыслях человека и оказывать влияние на них. Коммерческое применение устройств, которые собирают нейронные данные, открывает новые возможности для профилирования и слежки. Это имеет самое непосредственное отношение к свободе мысли. При том что такие устройства предлагают потенциально революционные решения, особенно в медицине, считается, что они могут открыть беспрецедентные возможности для проникновения во внутреннюю сферу, которая долгое время считалась неприкосновенной — неприкосновенность разума. При использовании в коммерческих целях эти технологии способны поддерживать и укреплять силу и изоционность влияния крупных онлайн-платформ на индивидуальный выбор каждого человека, в связи с чем возникают новые проблемы в области этики и прав человека⁴⁶.

Технологии на основе искусственного интеллекта преимущественно разрабатываются и внедряются частными компаниями, которые действуют в коммерческой системе координат, где получение прибыли необходимо для обеспечения устойчивости деятельности. В результате эти технологии в значительной степени формируются вне демократического контроля и традиционного «общественного договора» между государством и гражданами, несмотря на их существенное влияние на сферы, представляющие общественный интерес. Это несоответствие означает, что ныне действующие органы

46 См., например, Nita A. Farahany, *The Battle for Your Brain. Defending the Right to Think Freely in the Age of Neurotechnology* [Борьба за ваш мозг. Защита права на свободу мысли в эпоху нейротехнологий], (New York: St Martin's Griffin, 2023).

управления не могут надлежащим образом контролировать широкое влияние этих технологий на общество. Кроме того, это означает, что коммерческие приоритеты этих частных субъектов могут совпадать или не совпадать с более широкими целями общественного благополучия, создавая проблемы в процессе обеспечения того, чтобы разработка ИИ всегда служила общему благу.

С учетом значительного влияния ИИ на демократические институты, права человека и взаимодействие между людьми, необходимо ответить на вопросы о том, кому, на каких условиях, в соответствии с какими нормами и с какой целью следует разрабатывать и внедрять такие технологии. Помимо преобладающих технологических гигантов, в сфере искусственного интеллекта есть множество стартапов и других конкурентов⁴⁷. Отраслевые, и особенно этические, стандарты только формируются и на практике часто зависят от экономического веса или деловой культуры отдельных компаний, а также от условий, в которых они работают.

В условиях «революции в сфере ИИ» государства должны взять на себя обязательства по защите прав и основных свобод человека. В то же время им следует создавать равные условия в этой сфере, поощрять честную конкуренцию между всеми участниками и не допускать снижения стандартов в борьбе компаний за лидерство. Государства обязаны создавать и защищать благоприятную среду для обеспечения свободы мысли и всех других прав человека, тщательно отслеживая влияние крупнейших технологических компаний, для того чтобы гарантировать подотчетность и соблюдение общественных интересов⁴⁸.

47 Rashi Shrivastara (ed.), [Forbes 2025 AI 50 List - Top Artificial Intelligence Companies Ranked](#) [Список Forbes 2025 AI 50 — рейтинг ведущих компаний в области искусственного интеллекта], Forbes, 10 April 2025; Kate Brennan et al., [Artificial Power: 2025 Landscape Report](#) [Способности ИИ: обзор за 2025 год].

48 К другим важным сферам, которые затрагивают свободу мысли, но не рассматриваются здесь, относится использование ИИ в целях наблюдения, особенно в целях национальной безопасности или обороны, на границах и при миграционном контроле, во время содержания под стражей, в военных целях или в условиях авторитарных режимов. См. также [Digital technologies at borders: A threat to people on the move](#) [Цифровые технологии на границах: угроза для перемещающихся лиц], UN OHCHR, 9 October 2023.

3. Алгоритмическое мышление и его последствия для прав человека

ИИ, предвзятость и мышление

Результаты работы зависят от целей и возможностей системы ИИ, качества данных и способа внедрения. Отдельная система ИИ будет отражать как предвзятость и «слепые зоны» тех, кто участвовал в ее разработке, так и те, которые заложены в наборах данных, которые использовались для ее создания и обучения. Такая система может быть развернута для целей, отличных от тех, для которых она создавалась и тестировалась, при этом некоторые проблемы могут присутствовать в самих алгоритмах машинного обучения⁴⁹. Например, в областях, которые предусматривают социальное взаимодействие (банковское обслуживание, социальная защита, жилье, занятость), элементы предвзятости в наборах данных имеют особенно негативные последствия, поскольку системы ИИ способны распространять и потенциально усугублять такую предвзятость⁵⁰. Кроме того, системы ИИ зачастую отражают ценности и реалии Глобального Севера, поэтому могут предоставить искаженную информацию о других регионах. Такие последствия являются рисками с точки зрения прав человека, в том числе права на свободу мысли. В то же время важно продолжать усилия по борьбе с предвзятостью, в частности, путем поощрения разнообразия

49 [Bias in Algorithms: Artificial Intelligence and Discrimination](#), EU Agency for Fundamental Rights (FRA) [Предвзятость алгоритмов, искусственный интеллект и дискриминация, Агентство ЕС по основным правам], (Vienna, FRA, 2022).

50 Например, на алгоритмы прогнозирования преступности может влиять ряд факторов, в том числе, если набор данных основан на ранее зарегистрированных показателях. Эти показатели, в свою очередь, могут быть подвержены влиянию предвзятого отношения к конкретным меньшинствам, проживающим в конкретных районах, которые, таким образом, чаще фигурируют в наборах данных как районы с высоким уровнем преступности, чем это было бы в контексте реальных (зарегистрированных и незарегистрированных) показателей преступности. Таким образом, алгоритмические прогнозы могут быть предвзятыми по отношению к этим меньшинствам, а их районы проживания могут столкнуться с чрезмерным полицейским контролем. Подробное объяснение предвзятости в системах прогнозирования преступности с использованием ИИ см. в FRA, [Bias in Algorithms](#) [Предвзятость алгоритмов], р. 31. [Закон ЕС об ИИ](#) запрещает использование ряда систем прогнозирования преступности с использованием ИИ, если они основаны исключительно на профилировании или личных характеристиках, за исключением случаев, когда они используются для дополнения оценок, сделанных человеком на основе объективных, проверяемых фактов, непосредственно связанных с преступной деятельностью (см. также Главу 6 далее).

на протяжении всего жизненного цикла ИИ, устранения технических проблем, а в ряде случаев отказа от использования отдельных типов ИИ, если риски слишком высоки, в рамках усовершенствования мышления, защиты разнообразия и обеспечения доступа к разным точкам зрения.

«Цикл обратной связи возникает, когда предсказания, сделанные системой, влияют на данные, используемые для обновления этой системы. Это значит, что алгоритмы воздействуют на алгоритмы, поскольку их рекомендации и предсказания влияют на реальность на местах. Затем эта реальность становится основой для сбора данных, предназначенных для обновления алгоритмов, поэтому результаты, которые генерирует система, становятся входными данными для этой же системы. Таким образом, любая предвзятость алгоритмов может со временем усиливаться и усугубляться».

— EU Agency for Fundamental Rights, [Bias in Algorithms: Artificial Intelligence and Discrimination](#) [Предвзятость алгоритмов, искусственный интеллект и дискриминация], (Vienna, FRA, 2022), p. 8.

Когда принятие решений и взаимодействие между людьми заменяются системами искусственного интеллекта или формируется чрезмерная зависимость от таких систем, искажаются условия для принятия обоснованных решений в связи с существующими или сгенерированными системой элементами неравенства и неточностями, которые эта система в дальнейшем усиливает. В тех случаях, когда предвзятый контент в социальных сетях или профилирование с помощью ИИ оказывают влияние на политическую и другую общественную деятельность, чаще всего в первую очередь страдают лица с конкретными защищенными характеристиками или моделями поведения⁵¹, а в конечном итоге это влияет на все общество.

51 Mehnaz Rafi, [When AI plays favourites: How algorithmic bias shapes the hiring process](#) [Когда ИИ проявляет симпатии: как алгоритмическая предвзятость влияет на процесс найма] *The Conversation*, 14 October 2024; см. также [ICO considers uses of neurotechnology in employment in the UK](#) [ICO рассматривает использование нейротехнологий в сфере занятости в Великобритании], Withers, 31 July 2023.

Отбор данных с целью устранения, а не сокращения, предвзятости, учитывая сложность и масштаб наборов данных, представляется сложной задачей. Даже в случае принятия мер по устранению предвзятости из инструментов ИИ с лучшими намерениями ошибки в этом процессе сами по себе могут приводить к (дальнейшим) нарушениям прав человека. Например, социальные сети уже широко применяют ИИ для модерации — фильтрации, выделения, удаления или ограничения распространения незаконного и оскорбительного контента. Пытаясь разобраться в сложностях человеческого общения на этих платформах, ИИ часто не понимает контекст и оттенки значения (например, сатиру), что ограничивает его возможности по модерации общения между людьми в целом, не говоря уже о сложностях, связанных с предвзятостью⁵². Кроме того, значение часто зависит от контекста, общего понимания и языка, а также может меняться со временем. При модерации с помощью ИИ в социальных сетях маргинализированные сообщества могут непреднамеренно подвергаться цензуре за использование определенного языка, в частности, защищенной контрречи, противостоящей вредным и ненавистническим высказываниям или оспаривающей их⁵³. В то же время контроль со стороны ИИ может избежать контент, в котором используется «закодированный» язык. Такие проблемы более ярко проявляются применительно к контенту на других языках, кроме английского, поскольку компании вкладывают меньше ресурсов в модерацию и проверку фактов на таких языках⁵⁴. Усилия по выявлению предвзятости в данных, преимущественно на основе ИИ, должны предусматривать привлечение разносторонне образованных людей с самого начала жизненного цикла систем ИИ.

Вызывает беспокойство тот факт, что сегодня социальные сети пересматривают свою политику в отношении модерации контента и защиты, связанной с доверием и безопасностью, в области предотвращения ненависти, преследований и распространения

52 Представитель ОБСЕ по вопросам свободы СМИ. В фокусе – искусственный интеллект и свобода слова: рекомендации по разработке политики, 20 января 2022 г.

53 FRA, *Bias in Algorithms: Artificial Intelligence and Discrimination*. [АОП ЕС, Предвзятость в алгоритмах: искусственный интеллект и дискриминация].

54 *Content Moderation in a New Era for AI and Automation* [Модерация контента в новую эпоху ИИ и автоматизации], Meta Oversight Board, September 2024; см. также Представитель ОБСЕ по вопросам свободы СМИ. В фокусе – искусственный интеллект и свобода слова: рекомендации по разработке политики.

недостовверного контента. Они существенно сократили число сотрудников, которые обеспечивали эффективную модерацию⁵⁵, ухудшив свою работу по борьбе с вредоносным контентом⁵⁶. Такой контент наносит маргинализированным сообществам значительный ущерб (из-за выражения ненависти в сети они могут столкнуться с угрозами в реальной жизни), поэтому меры по искоренению предвзятости необходимо усилить⁵⁷.

Одной из важных особенностей цифровой среды является доминирование английского языка. Такая ситуация не отражает разнообразие языков, на которых говорят в мире, и усугубляет цифровое неравенство с точки зрения доступа к Интернету для людей, которые не говорят на английском или других распространенных языках⁵⁸. Генеративный ИИ способствует усугублению этой проблемы, поскольку он обучается на ограниченном количестве языков и, следовательно, наполняет информационную среду преимущественно контентом на этих языках. Эта тенденция угрожает языковому разнообразию и, по-видимому, не изменится без значительных инвестиций со стороны компаний⁵⁹.

Родной язык важен не только для культуры и идентичности: он играет существенную роль в том, как люди организуют свои мысли в *forum internum* (внутреннее измерение убеждений и взглядов человека) и как формируется их мировоззрение. Позитивным достижением ИИ в этой

55 David Evan Harris and Aaron Shull, [Generative AI, Democracy and Human Rights](#), Center for International Governance Innovation, Policy Brief No. 12, February 2025 [Генеративный искусственный интеллект, демократия и права человека, Центр инноваций в области международного управления].

56 Joel Kaplan, [More Speech and Fewer Mistakes](#) [Больше свободы слова и меньше ошибок], Meta, 7 January 2025; Clare Duffy, [Calling women 'household objects' now permitted on Facebook after Meta updates its guidelines](#) [Теперь в Facebook можно называть женщин «предметами домашнего обихода» после обновления правил Meta], *cnn.com*, 8 January 2025; Dia Kayyali, [Meta's Content Moderation Changes are Going to Have a Real World Impact. It's Not Going to be Good](#) [Изменения в модерации контента Meta окажут реальное влияние на мир. Оно не будет хорошим], *TechPolicy.press*, 9 January 2025; Adrian Kopps, [Four key policy changes of X under Musk](#) [Четыре ключевых изменения политики X при Маске], Digital Society Blog, Alexander von Humboldt Institut für Internet und Gesellschaft, 28 October 2024.

57 [Язык ненависти](#), веб-страница ООН; [X's design and policy choices created fertile ground for inflammatory, racist narratives targeting Muslims and migrants following Southport attack](#) [Дизайн и политические решения X — благодатная почва для подстрекательских расистских нарративов, направленных против мусульман и мигрантов после нападения в Саутпорте], Amnesty International, 6 August 2025.

58 [Internet Access: UNESCO and ICANN join forces to improve linguistic diversity online](#) [Доступ в Интернет: ЮНЕСКО и ICANN объединяют усилия для улучшения языкового разнообразия в Интернете], UNESCO, 27 February 2025.

59 Viorica Marian, [AI could extinguish languages and ways of thinking](#) [Искусственный интеллект может уничтожить языки и способы мышления], *The Washington Post*, 19 April 2023.

области — наряду с другими примерами вклада ИИ в общественное благо⁶⁰ — стало его успешное использование для документирования и сохранения исчезающих языков, а также возрождения утраченных⁶¹. Ключевым здесь является наличие (или отсутствие) этики и понимания прав человека, которые влияют на разработку, внедрение и использование ИИ.

Без надлежащего обеспечения разнообразия и представленности, в частности, путем преодоления цифрового неравенства, будущее систем ИИ, при масштабном использовании которых воспроизводятся многочисленные дисбалансы власти, может привести к исчезновению мировоззрения, не представленного в системах ИИ.

Человеческий надзор, прозрачность и подотчетность

С точки зрения обеспечения соблюдения прав человека особенно важное значение имеет уровень вмешательства человека в процесс принятия решений. Полностью автоматизированное принятие решений не предполагает участия человека (в том числе применения человеческой интуиции и оценочных суждений) в процессе формирования системой ИИ конечного результата. Кроме того, сложность систем генеративного ИИ затрудняет и может сделать невозможным понимание или объяснение того, как был получен тот или иной результат. Еще одна трудность во взаимодействии человека и машины — это «автоматизационная предвзятость». По данным исследований, человек, как правило, больше доверяет машинам, чем людям, и, следовательно, с большей готовностью воспринимает результаты, сгенерированные системой ИИ⁶². Свобода мысли да и само значимое участие человека подрываются в ситуации, когда становится сложнее проверять и ставить под сомнение сгенерированные результаты, особенно в

60 См. ООН [AI for Good](#) [ИИ во благо] и Stéphanie Bascou, « [IA d'intérêt public](#) » : cette ONG veut créer un label éthique, similaire au « [Commerce équitable](#) », et facilement identifiable pour les consommateurs [ИИ в интересах общества: эта НПО хочет создать этическую маркировку, аналогичную «справедливой торговле» и легко идентифицируемую для потребителей], *01.net*, 5 October 2024.

61 [AI: The Unexpected Hero in the Battle to Save Dying Languages](#) [Искусственный интеллект: неожиданный герой в битве за спасение вымирающих языков], *apolitical website*, 20 October 2024.

62 Chris Baraniuk, [Why we place too much trust in machines](#) [Почему мы слишком доверяем машинам], *BBC*, 20 October 2021.

ответственных ситуациях, даже если человек участвует в принятии решений о том, как будут применяться эти результаты или какие действия будут предприниматься на их основании⁶³. Степень риска зависит от области применения ИИ. Например, врачи должны полностью понимать систему ИИ, которую они используют, и уметь преодолевать предвзятость автоматизации, систематически оценивая сгенерированные системой ИИ результаты по каждому пациенту из-за очень высокого риска негативных последствий для здоровья⁶⁴.

Зачастую возможность проверить и оспорить результаты еще больше затруднена в связи утверждениями о праве интеллектуальной собственности в отношении кода, параметров и/или иных внутренних механизмов ИИ. Эти аргументы используются для того, чтобы не раскрывать способ получения того или иного результата. В отсутствие прозрачности и возможности контроля над информационной средой (например, того, какой контент рекомендательные системы социальных платформ отображают в пользовательских лентах) принятие обоснованных решений и формирование реального понимания своего окружения еще больше затрудняется.

«Понимаемость — это то, насколько любая система может быть понятна любому человеку. Если предсказуемость системы связана с вопросом «*Что будет делать система?*», то понятность связана с вопросом «*Почему она это делает?*»

— Arthur Holland Michel, *The Black Box, Unlocked* [Черный ящик: разгадка], UNIDIR, 22 September 2020, p. 9.

Тот факт, что понять, как ИИ принимает решения, затрагивающие права человека, не представляется возможным (из-за чрезмерной сложности системы или потому, что информация о ее работе и средствах ее контроля отсутствует в открытом доступе) вызывает серьезные вопросы

63 Подробнее о предсказуемости и понятности см. Arthur Holland Michel, *The Black Box, Unlocked* [Черный ящик: разгадка], UNIDIR, 22 September 2020.

64 *The application of artificial intelligence in healthcare and its impact on the “patient-doctor” relationship* [Применение искусственного интеллекта в здравоохранении и его влияние на отношения между пациентом и врачом], Council of Europe, Steering Committee for Human Rights in the Fields of Biomedicine and Health (CD-BIO), September 2024, p. 20.

по поводу подотчетности и прозрачности, двух главных принципов надлежащего функционирования систем правосудия и демократии.

«1. Цифровые технологии преобразуют наш мир кардинальным образом. Они дают огромные потенциальные преимущества для обеспечения благополучия и развития людей и обществ и для нашей планеты. (...) 3. Мы признаем, что темпы развития и мощь новейших технологий не только открывают новые возможности для человечества, но и создают новые риски, и о некоторых из них еще не все известно. Мы признаем необходимость выявления и сокращения рисков и обеспечения контроля за технологиями со стороны человека, который осуществляется таким образом, чтобы способствовать устойчивому развитию и полному осуществлению прав человека».

— **Глобальный цифровой договор**, приложение к Резолюции A/RES/79/1 «Пакт во имя будущего», принятой Генеральной Ассамблеей ООН 22 сентября 2024 г.

Для минимизации этих рисков нужен другой подход⁶⁵ к разработке ИИ — ответственный и ориентированный на права человека, обеспечивающий регуляторный надзор и надлежащую понятность, прозрачность и подотчетность. Кроме того, значимый человеческий контроль имеет решающее значение для того, чтобы системы ИИ помогали обеспечивать наилучшие для людей результаты и не нарушали при этом их права, в том числе право на свободу мысли.

65 [Recommendation on the Ethics of Artificial Intelligence](#) [Рекомендация по вопросу об этике искусственного интеллекта], UNESCO, 23 November 2024; см. также Rumman Chowdhury and Dhanya Lakshmi, [“Your opinion doesn’t matter, anyway”: exposing technology-facilitated gender-based violence in an era of generative AI](#) [Ваше мнение все равно не имеет значения]: разоблачение гендерного насилия с применением технологий в эпоху генеративного ИИ], (Paris: UNESCO 2023); или раздел [Ресурсы](#) на сайте организации Humane Intelligence.

4. Свобода мысли в изменившемся гражданском пространстве

Как новая архитектура гражданского пространства влияет на свободу мысли

Подлинная свобода мнений и их выражения зависит от подлинной свободы мысли⁶⁶. Как говорилось выше, мысль и мнение являются частью *forum internum* — частной внутренней сферы каждого человека. Это пространство, которое имеет абсолютную защиту от вмешательства в соответствии с международным правом в области прав человека. Согласно демократической теории свободная мысль формируется и развивается в ходе разговора, когда люди свободно обмениваются мнениями и делятся информацией⁶⁷. В таком идеальном гражданском пространстве должна существовать возможность выражать, оспаривать и проверять все идеи и мнения, чтобы обеспечить в конечном итоге возобладание наилучших идей и истин. С возможностью такого динамичного обмена, при котором идеи могут сближаться и противостоять друг другу, связана мощная защита и поддержка свободы слова, т. е. выражения мыслей и мнений. Это имеет критическое значение для здоровой общественной дискуссии, необходимой при демократии. Общественно значимая журналистика играет ключевую роль в обеспечении демократической дискуссии, основанной на фактах, а также, среди прочего, свободы ассоциации и права на организацию. В то же время существуют и ограничения. Например, в соответствии с международным правом в области прав человека государства должны бороться с риторикой ненависти, если она выходит за определенные рамки⁶⁸.

66 Sjors Ligthard, Christoph Bublitz, Thomas Douglas, Lisa Forsberg and Gerben Meynem, [Rethinking the Right to Freedom of Thought: A Multidisciplinary Analysis](#) [Переосмысление права на свободу мысли: междисциплинарный анализ] in *Human Rights Law Review*, 2022, 22, pp. 1-14.

67 См. также Simon McCarthy-Jones, *Freethinking: Protecting Freedom of Thought Amidst the New Battle for the Mind*, [Свободомыслие: защита свободы мысли на фоне новой битвы за разум] (London: Oneworld, 2023).

68 МПГПП, Статья 20:

“1. Всякая пропаганда войны должна быть запрещена законом.

На фоне развития крупных технологических корпораций коммуникация переместилась в социальные сети, которые сегодня обеспечивают взаимодействие между людьми и выступают в роли контролеров новостных и информационных потоков. Эти платформы используют мощные алгоритмы машинного обучения для реорганизации контента такими способами, которые имеют последствия для свободы мысли. Они кардинальным образом меняют то, как распространяется информация, мнения и идеи в отличие от информационной среды, в значительной степени сформированной традиционными (общественно значимыми) СМИ⁶⁹.

Рекомендательные системы на основе алгоритмов в Интернете создаются в первую очередь исходя из коммерческих соображений. Такие системы оптимизированы с целью достижения максимальной вовлеченности пользователей или роста продаж, они не ориентируются на достоверность, разнообразие или общественный интерес, в том числе на плюрализм медиа и общественно значимую журналистику⁷⁰. Максимизация вовлеченности и стимулирование покупок предполагают использование моделей поведения в прошлом и индивидуального профилирования для предсказания того, с каким контентом пользователи будут взаимодействовать с наибольшей вероятностью. В результате контроль пользователей над тем, что они могут видеть, ограничивается, даже если они сохраняют автономию и могут выбирать, хотят ли они взаимодействовать с определенным контентом или нет. При этом они могут не осознавать, что в связи с персонализацией на основе алгоритмов контент, который они видят, отличается от того, что видят другие люди, в том числе из их ближайшего окружения.

2. Всякое выступление в пользу национальной, расовой или религиозной ненависти, представляющее собой подстрекательство к дискриминации, вражде или насилию, должно быть запрещено законом.”

69 «В отличие от традиционных алгоритмов, которые задают разработчики, алгоритмы машинного обучения «учатся» на входных данных, выявляя в них корреляции. Обученный алгоритм, или модель машинного обучения, в дальнейшем может автоматизировать решения. Например, алгоритм, обученный на данных о кликах по рекламным объявлениям, способен выяснить, что женщины чаще кликают по рекламе леггинсов для занятий йогой, чем мужчины. Такая модель будет показывать больше соответствующих объявлений женщинам». См. Karen Hao, [The Facebook whistleblower says its algorithms are dangerous. Here's why](#) [Разоблачитель из Facebook утверждает, что алгоритмы компании опасны. Вот почему], *MIT Technology Review*, 5 October 2021.

70 Представитель ОБСЕ по вопросу о свободе СМИ. В фокусе - искусственный интеллект и свобода слова: рекомендации по разработке политики.

Такая модель позволяет нескольким платформам оказывать существенное влияние на информационную экосистему каждого человека. Она способствует фрагментации новостного медиаландшафта, ослаблению независимой журналистики и потенциальному подрыву общего понимания реальности, которое долгое время определяло взаимодействие между людьми, их выбор и участие в демократических процессах⁷¹.

«Методы, с помощью которых онлайн-платформы курируют контент при помощи рекомендательных систем, недостаточно прозрачны и очень редко подвергаются общественному и/или государственному контролю. <...> Рекомендательные системы контента также могут иметь непредвиденные последствия для масштабных общественных целей и способны негативно влиять на абсолютное право на свободу мысли и мнения. Кроме того, процессы рекомендательных систем интернет-посредников, как правило, исключают возможность выбора, контроля и автономность отдельных пользователей, которые являются необходимыми условиями для обеспечения автономии личности в процессе поиска и передачи разнообразной информации и идей.<...> Существует множество доказательств того, что манипулятивный потенциал онлайн-платформ способен направлять и усиливать определенные общественные нарративы и типы дискурса по сравнению с другими. Для стран с нестабильными или деспотичными политическими

71 Представитель ОБСЕ по свободе СМИ. В фокусе - искусственный интеллект и свобода слова: [рекомендации по разработке политики](#); в более широком плане работа Представителя ОБСЕ по свободе СМИ на тему «СМИ и крупные технологические компании»; см. также [Декларацию Комитета Министров о манипулятивных возможностях алгоритмических процессов](#), принятую Комитетом Министров Совета Европы 13 февраля 2019 года (на англ. языке): «8. Современные инструменты машинного обучения развивают способность не только прогнозировать выбор, но и влиять на эмоции и мысли, а также изменять предполагаемый ход событий, иногда на уровне подсознания. <...> 9. Тонкие, подсознательные и персонализированные уровни алгоритмического воздействия могут оказывать существенное влияние на когнитивную автономию личности и ее право на формирование мнения и принятие независимых решений. <...> Центральные принципы [Совета Европы] – права человека, демократия и верховенство права – основываются на фундаментальной вере в равенство и достоинство всех людей как независимых моральных субъектов. <...> Комитет Министров обращает внимание на растущую угрозу для права человека формировать мнения и принимать решения независимо от автоматизированных систем, которая исходит от передовых цифровых технологий. Особое внимание следует уделить их способности использовать персональные и неперсональные данные для сортировки и микротаргетинга людей, выявления индивидуальных уязвимостей и использования точных прогнозных знаний, а также для перестройки социальной среды в целях достижения конкретных целей и преследования корыстных интересов».

системами эта сила убеждения, усиленная алгоритмами, может иметь катастрофические последствия в плане реализации индивидуальных прав человека.

— Представитель ОБСЕ по вопросам свободы СМИ. **В фокусе – искусственный интеллект и свобода слова: рекомендации по разработке политики**, 20 января 2022 г. С. 77–93.

Представитель ОБСЕ по вопросам свободы СМИ провел масштабную работу по изучению и разъяснению влияния ИИ на цифровой рынок идей, общественно значимые средства массовой информации, демократические процессы, свободу мнений и их выражения, а также предложил всеобъемлющие рекомендации для государств-участников.

См. дополнительно Представитель ОБСЕ по вопросам свободы СМИ. **В фокусе – искусственный интеллект и свобода слова. SAIFE Resource Hub.**

По мере развития возможностей ИИ повышается эффективность рекламы на базе бизнес-модели, в основе которой лежит сбор данных с использованием ИИ, анализирующей индивидуальные и коллективные формы поведения и характеристики. Такая система отслеживает активность пользователей на различных сайтах и устройствах и предлагает персонализированный контент, включая новостной, с целью максимизации вовлеченности и стимулирования продаж.

С развитием социальных сетей появились новые виды контактов и коммуникации, которые без сомнения обеспечили возможность мобилизации общественного мнения по вопросам, вызывающим всеобщую озабоченность, и позволили выражать свое мнение тем, кто подвергался цензуре. В то же время, в отличие от традиционных общественных пространств, эти платформы меняют способ взаимодействия людей как друг с другом, так и с группами. Такое коммуникационное пространство предполагает использование психологических предубеждений человека, особенно в отношении «своих» и «чужих», а также потребность людей в принадлежности к какой-либо группе, для того чтобы предложить персонализированный контент в соответствии с критериями, которые, как правило, не являются публичными (т. е. речь идет о том, как настроены алгоритмы, управляющие контентом, и какую цель они преследуют). Соцсети

способствовали трансформации значительной части общественного пространства и способов нашего взаимодействия с ним, поддерживая при этом полную иллюзию наличия выбора⁷². Благодаря тонкому, но систематическому воздействию люди становятся уязвимыми к влиянию и манипулированию их мыслями и поведением, что подрывает идею самостоятельного мышления и снижает нашу субъектность⁷³.

Доминирование крупных технологических компаний, по сути, не оставляет человеку иного выбора, кроме как принять надзорную модель и использовать социальные сети. Для многих повседневных социальных и профессиональных контактов это стало необходимостью, особенно в случае, если групповые мероприятия или другая важная информация публикуются на определенной платформе. Это оказывает влияние на свободу мысли. Поскольку люди подчиняются алгоритмическим системам компаний в плане фильтрации значительной части своего восприятия мира, они рискуют утратить контроль над своей индивидуальностью. Даже само понимание того, что за ними следят, может влиять на то, что люди думают и как они действуют, опасаясь

72 В контексте политической поляризации часто упоминается возникновение в новой информационной среде эхо-камер как результата склонности людей — а затем и реальной возможности — взаимодействовать в социальных сетях с теми, кто похож на них и чьи взгляды они разделяют. В литературе нет единого мнения по поводу эхо-камер, при этом соответствующие исследования ограничены в географическом и лингвистическом отношении, и, как правило, не опираются на данные, необходимые для надлежащего анализа. По данным масштабного совместного метаисследования политической активности, есть доказательства того, что в социальных сетях люди вступают в однородные партийные объединения. Вместе с тем ситуация намного сложнее. Авторы также подчеркивают, что, хотя социальные сети предоставляют человеку более широкий доступ к новостям, информации и идеям, чем он получает в офлайн-пространстве, разные мнения и информация в интернете воспринимаются с поляризованных позиций. Реагируя на противоположное мнение, люди не участвуют в демократической дискуссии, а ищут одобрение в онлайн-сообществе, противопоставляя «своих» и «чужих». В конечном итоге это приводит к еще более глубокому укоренению точек зрения и дальнейшей поляризации. См. Jonathan Haidt and Chris Bail (ongoing), *Social Media and Political Dysfunction, A collaborative review* [Социальные сети и политическая дисфункция, совместный обзор, продолжающееся исследование], New York University, Chapter 2.4.

73 Шошана Зубофф, которая предложила термин «надзорный капитализм», описывает его как новую экономическую логику, в основе которой лежит получение персональных данных с целью формирования поведения пользователей. <...> «надзорные капиталисты обнаружили, что наиболее ценные с точки зрения предсказания поведенческие данные поступают в том случае, если вмешаться в ход игры – подтолкнуть, уговорить, подстроить и направить поведение в сторону, которая принесет прибыль. Конкурентное давление привело к сдвигам, в результате которых автоматизированные машинные процессы не только *познают* наше поведение, но и в широких масштабах *формируют* наше поведение. С этой переориентацией со знания на силу уже недостаточно автоматизировать *потоки информации о нас*; теперь цель состоит в том, чтобы *автоматизировать нас*». Шошана Зубофф. Эпоха надзорного капитализма. Битва за человеческое будущее на новых рубежах власти, Издательство Института Гайдара, 2022 г. С. 14.

возможных последствий⁷⁴. Следует отметить, что ЕС начал принимать меры по решению проблемы таргетированной рекламы в различных аспектах своей законодательной базы. На данный момент рассмотрен ряд дел, а некоторые иски находятся в процессе рассмотрения⁷⁵. Некоторые из них окажут влияние на защиту автономии и субъектности, а значит и свободы мысли, в новом гражданском пространстве.

«...тот факт, что оператор социальной сети занимает доминирующее положение на рынке социальных сетей, сам по себе не исключает возможности для пользователей этой сети дать действительное согласие <...> на обработку их персональных данных этим оператором. В то же время это важный фактор при определении того, было ли такое согласие действительным и в особенности добровольным, что должен доказать сам оператор».

— [Пресс-релиз № 113/23](#), Суд Европейского союза, решение по делу C-252/2001 | *Meta Platforms и другие* (Общие условия использования социальных сетей), Люксембург, 4 июля 2023 г., п. 154. (на англ. языке)

Более того, в академической литературе на тему свободы мысли предполагается, что новая среда оказывает негативное влияние на ментальную автономию человека из-за: 1) слежки и использования персональных данных, содержащих очень подробную информацию об отдельных лицах или группах; 2) непрозрачности технологий, как в отношении использования ИИ для генерации и распространения контента,

74 После разоблачений Сноудена о слежке со стороны правительства люди изменили свое поведение в Интернете и исключили из поисковых запросов темы, которые, на их взгляд, могли бы привлечь к ним внимание государства. См. [Jon Penney, Chilling Effects: Online Surveillance and Wikipedia Use](#) [Негативные последствия: слежка за пользователями и Википедия], *Berkley Technology Law Journal*, Vol. 31, No. 1, pp. 117-183, 27 April 2016, См. также Ахмед Шахид, [Доклад Специального докладчика](#), п. 54.

75 [Commission finds Apple and Meta in breach of the Digital Markets Act](#) [Европейская комиссия постановила, что Apple и Meta нарушили закон «О цифровых рынках»], пресс-релиз Европейской комиссии от 23 апреля 2025 г.; см. также [пресс-релиз № 116/24](#), решение суда по делу C-446/21 | *Шремс* (Передача данных широкой общественности), Суд Европейского союза, Люксембург, 4 октября 2024 г.; и [пресс-релиз № 113/23](#), решение суда по делу C-252/2001 | *Meta Platforms и другие* (Общие условия использования социальной сети), Суд Европейского союза, Люксембург, 4 июля 2023 г. (на англ. языке). Аналогичный иск был рассмотрен в Великобритании в начале 2025 года. См. Dan Milmo, [Meta to stop targeting UK citizen with personalised ads after settling privacy case](#), [Meta прекратит показывать персонализированную рекламу гражданам Великобритании после урегулирования иска о нарушении конфиденциальности] *The Guardian*, 22 March 2025.

так и процессах работы самой технологии; 3) эксплуатации когнитивных искажений; 4) ее беспрецедентного охвата, формирующего общественные предпочтения и выбор в обход демократических процессов⁷⁶.

Как «информационная путаница» меняет реальный выбор

Реорганизация нашей информационной экосистемы повлияла на то, как происходит демократическая дискуссия и как осуществляется выбор, а также серьезно затронула общественно значимые СМИ. Широкое распространение цифровых систем ИИ, которые благодаря цифровым платформам распространяют гораздо быстрее, чем раньше циклы ложных сведений и дезинформации для достижения определенных политических, идеологических или коммерческих целей⁷⁷, привело к «информационной путанице»⁷⁸. Это явление подорвало доверие к независимым СМИ и к достоверности информации, а также ведет к размыванию истины. На фоне растущей неопределенности снижается доверие общества к институтам и информационной экосистеме в целом, что, в свою очередь, влияет на качество процесса принятия обоснованных решений, необходимых как для участия в демократических процессах, так и для повседневной жизни⁷⁹.

76 Nina Keese and Mark R. Leiser, Online Manipulation as a Potential Interference with the Right to Freedom of Thought [Манипуляции в Интернете как потенциальное нарушение права на свободу мысли] in Patrick O’Callaghan and Bethany Shiner (eds.), *The Cambridge Handbook of the Right to Freedom of Thought*, (Cambridge: Cambridge University Press, 2025); см. также Kate Brennan et al, **Artificial Power: 2025 Landscape Report** [Способности ИИ: обзор за 2025 год].

77 Дезинформация и свобода мнений и их свободное выражение. Доклад Специального докладчика ООН по вопросу о поощрении и защите права на свободу мнений и их свободное выражение Айрин Хан, 13 апреля 2021 г., **A/HRC/47/25**, п. 2.

78 См. Claire Wardle PhD and Hossein Derakhshan, Information Disorder: Towards an interdisciplinary framework for research and policy making [Информационная путаница: на пути к междисциплинарным рамкам для исследований и разработки политики], (Strasbourg: Council of Europe, 2018) Уордл и Деракшан для описания информационной путаницы в новой коммуникационной среде используют концептуальную основу, в рамках которой выделяют три типа (дезинформация, ложная информация и вредная информация); три фазы (создание, производство и распространение), а также элементы информационной путаницы (агент, сообщение и интерпретатор). Они рассматривают вред и недостоверность информации для того, чтобы определить ее типы, составляющие экосистему «информационной путаницы». Так, ложная информация — это недостоверная информация, при распространении которой не предполагается причинение вреда. Дезинформация — это ложная информация, сознательно распространяемая, чтобы причинить вред. Вредная информация — это достоверная информация, которая распространяется с намерением причинить вред (например, публикация персональных данных).

79 **Дезинформация и свобода мнений и их свободное выражение**. Доклад Специального докладчика ООН по вопросу о поощрении и защите права на свободу мнений и их свободное выражение Айрин Хан.

По данным Представителя ОБСЕ по вопросам свободы СМИ⁸⁰, генеративный ИИ создает дополнительные угрозы для журналистики. Он заменяет оригинальный контент, который создают журналисты, а также способствует возникновению проблем, связанных с плагиатом и выплатой вознаграждения. Все это может способствовать дальнейшему вытеснению журналистики и контента, которые представляют интерес для общества, из публичной сферы⁸¹.

Одновременно с изменением роли и возможности СМИ действовать в общественных интересах, поддерживая здоровую информационную среду⁸², генеративный интеллект и возможность создания дипфейков, которые становится все труднее обнаружить, создают дополнительные риски для информационного поля и особенно для избирательных процессов.

Такие технологии легко использовать для манипуляций, распространения дезинформации и ложных сведений, которые приводят к разрушительным последствиям. Дипфейки широко используются при гендерных домогательствах, в частности, для создания и распространения изображений сексуального характера. В первую очередь их жертвами становятся женщины, девочки и ЛГБТИ. С особыми рисками сталкиваются женщины в политике, а также женщины, которые живут в консервативных патриархальных обществах, где дипфейки, изображающие человека в «неподобающей» ситуации, могут привести к серьезным последствиям⁸³.

Например, установлено, что злоумышленники могут использовать модели генеративного ИИ с открытым исходным кодом для совершения

80 См. [Отчет по итогам семинара «Технологические корпорации и свобода СМИ»](#), Представитель ОБСЕ по вопросам свободы СМИ, 15 октября 2024 г. (на англ. языке).

81 Помимо журналистики, с генеративным интеллектом связаны вопросы соблюдения Статьи 15 МПЭСКП, которая касается права каждого человека на защиту моральных и материальных интересов, вытекающих из любой научной, литературной или художественной продукции, автором которой он является.

82 См. работа Представителя ОБСЕ по вопросам свободы СМИ в рамках инициативы [«СМИ и крупные технологические компании»](#) | ОБСЕ. Кроме того, в 2025 году Представитель ОБСЕ по вопросам свободы СМИ работает над руководством для государств по защите свободы СМИ в эпоху крупных технологических компаний и искусственного интеллекта в рамках поддержки развития здоровых информационных онлайн-пространств.

83 Наблюдательный совет Meta, веб-страница [«Модерация контента в новую эпоху ИИ и автоматизации»](#) (на англ. языке); см. также проект БДИПЧ ОБСЕ [«CHANGE: использование мандата в области человеческого измерения для поощрения гендерного равенства»](#) (на англ. языке).

или распространения гендерного насилия в цифровом пространстве⁸⁴. Несмотря на то что генеративный ИИ не является источником гендерного насилия, в отсутствие регулирования его способность генерировать, воспроизводить и распространять гендерное насилие и другие виды вредного контента в больших масштабах и с минимальными затратами делает его более разрушительным в руках таких пользователей. При этом пострадавшие вытесняются из онлайн-пространства и/или оказываются слишком запуганными, чтобы реагировать на такие действия⁸⁵.

Технической особенностью моделей генеративного интеллекта являются «галлюцинации». Это контент, сгенерированный ИИ, который не соответствует фактам или является полностью вымышленным, но, как правило, очень правдоподобным, вследствие чего его трудно, а иногда и невозможно, выявить. Такое явление возникает из-за особенностей работы генеративного ИИ: это обусловлено не только проблемами с базой данных, с которой работает ИИ (в частности, неточности или ложная информация, например, в Интернете), но и с тем, как сконструирован генеративный ИИ. Предполагается, что он должен «сотрудничать» и выдавать ответ⁸⁶. Как признала компания OpenAI, галлюцинации наряду с такими элементами, как преднамеренная дезинформация и социальные предубеждения, могут поставить под сомнение всю информационную среду и подорвать нашу способность отличать факт от вымысла⁸⁷. Такая «потеря чувства реальности» оказывает значимое влияние как на субъектность и автономию человека так и на свободу мысли. Компании и исследователи работают над уменьшением галлюцинаций⁸⁸, в том числе с помощью технических решений, которые позволяют увидеть источники, что, по многим причинам, имеет важное значение. Вместе с тем маловероятно, что

84 Chowdhury and Lakshmi, “Your opinion doesn’t matter, anyway” [«Ваше мнение все равно не имеет значения»].

85 Laura Bates, [Online brothels, sex robots, simulated rape: AI is ushering in a new age of violence against women](#) [Онлайн-бордели, секс-роботы, имитация изнасилования: ИИ открывает новую эру насилия в отношении женщин], *The Guardian*, 3 June 2025.

86 Kate Brennan et al, [Artificial Power: 2025 Landscape Report](#), p. 49 [Способности ИИ: обзор за 2025 год].

87 Colleen M Shannon, P. Eng., LL.M., [Do AI Hallucinations Disguise Gender Bias?](#), [Галлюцинации в ИИ маскируют гендерные предрассудки?], *Medium*, 27 September 2023.

88 Reese Rogers, [Reduce AI Hallucinations with This Neat Software Trick](#) [Как уменьшить галлюцинации ИИ с помощью одной программной хитрости], *Wired*, 14 June 2024

эту проблему можно решить в рамках существующих моделей, не пожертвовав при этом производительностью генеративного ИИ⁸⁹.

Закон ЕС об ИИ — это первый нормативный акт по регулированию использования генеративного искусственного интеллекта, включая продвинутые дипфейки, направленный на защиту людей от злоупотреблений с применением ИИ. Несмотря на то что одни группы пострадали больше других, «информационная путаница» угрожает свободе мысли, правам человека и общественной безопасности на глобальном уровне⁹⁰.

Как ИИ влияет на связи между людьми, развитие мышления и право на забвение

Способы восприятия и понимания человеком окружающего мира, в том числе благодаря критическому мышлению, формируются на базе их образования, воспитания и условий развития мозга по мере взросления. Искусственный интеллект (особенно генеративные модели) оказал серьезное влияние на образование и продолжит влиять на него в будущем. Вероятно, он будет изменять процессы преподавания и обучения, причем настолько стремительно, что возможностей для содержательного обсуждения плюсов и минусов и для принятия обоснованных решений практически не останется. Цифровая и медийная грамотность необходимы для всех, но учебные программы следует обновить с тем, чтобы предоставить учащимся всех возрастов знания и навыки для того, чтобы понимать среду, которую все больше определяет ИИ, ориентироваться в ней и осуществлять над ней контроль.

Применение ИИ в сфере образования имеет ряд потенциальных преимуществ. В то же время его использование в этой области, как и использование генеративного ИИ, связано с рисками, которые необходимо учитывать. ЮНЕСКО обращает внимание на потенциальные этические проблемы, связанные с внедрением генеративного ИИ в образовательных учреждениях, и призывает

89 Kate Brennan et al, *Artificial Power: 2025 Landscape Report*, p. 49 [Способности ИИ: обзор за 2025 год].

90 См. также Marina Nord et.al., *Media Freedom, Democracy, and Security* [Свобода СМИ, демократия и безопасность].

провести комплексную долгосрочную оценку воздействия⁹¹. Основное беспокойство вызывает потенциальное снижение взаимодействия между людьми. Такое взаимодействие и общение лежат в основе социально-эмоциональных аспектов обучения, которые важны для когнитивного развития и эмоционального благополучия.

Существует риск того, что ответы или результаты, которые предлагаются системами генеративного ИИ, заменят такой опыт, как обучение и приобретение навыков мышления. По данным научных исследований, использование генеративного ИИ снижает способность к критическому мышлению⁹². Последние исследования, проведенные компаниями, также свидетельствуют о деградации навыков мышления более высокого порядка среди специалистов, использующих в своей работе модели генеративного ИИ⁹³. Кроме того, есть опасения, что генеративный ИИ предоставляет пользователям ограниченный спектр мнений, поскольку, как правило, воспроизводит «типовые» ответы, что приводит к дальнейшей маргинализации недостаточно представленных мнений⁹⁴.

Для того чтобы ИИ играл позитивную роль в образовательной сфере, необходимо знать, каким образом он влияет на обучение и приобретение навыков мышления при интеграции в образовательную среду, а затем обеспечить возможности для обучения специалистов в сфере образования и предоставить им четкие рекомендации. Чтобы получить положительные результаты, системы ИИ следует использовать в соответствии с образовательными целями, связанными с надлежащим развитием мыслительных способностей и благополучием человека. Например, с его помощью можно развивать персонализированное обучение на адаптивных образовательных платформах, обеспечивать доступ к образовательным ресурсам в районах, где такой доступ

91 Miao and Holmes, [Guidance for generative AI in education and research](#) [Руководство по генеративному ИИ в образовании и исследовательской деятельности].

92 Andrew R Chow, [ChatGPT's Impact on Our Brains According to an MIT Study](#) [Влияние ChatGPT на наш мозг по данным исследования MIT], *Time*, 23 June 2025

93 [Despite plans to invest \\$80 bn, Microsoft admits that AI is making us dumb](#) [Несмотря на свои планы по инвестированию 80 миллиардов долларов, Microsoft признает, что ИИ делает нас глупее], *Business Standard*, 17 February 2025, источник Parmy Olson, Bloomberg; [Amanda Silberling, Is AI making us dumb?](#) [ИИ делает нас глупее?] *TechCrunch*, 10 February 2025.

94 Miao and Holmes, [Guidance for generative AI in education and research](#) [Руководство по генеративному ИИ в образовании и исследовательской деятельности].

затруднен, а также помогать учителям путем выполнения повседневных задач, например, автоматической проверки тестов с множественным выбором ответа. Таким образом, использование ИИ будет создавать пространство и возможности для развития свободы мысли.

Вызывает тревогу использование в новых генеративных ИИ имитирующих человеческое общение чат-ботов в качестве друзей или спутников. В этой связи особому риску подвергаются молодые люди, а также люди с проблемами психического здоровья. Кроме того, что они могут усугубить ментальные проблемы, чат-боты способны поощрять нездоровое поведение и даже довести человека до самоубийства⁹⁵. Особую обеспокоенность вызывает использование ИИ для решения проблем в области психического здоровья: разработка соответствующих программ помощи на основе ИИ является сложной задачей, при этом использование чат-ботов, в том числе предназначенных для других целей (например, для развлечений или маркетинга), связано с высокими рисками. Специалисты и профессиональные объединения психологов предупреждают о таких опасностях⁹⁶. Чтобы предотвратить вред и обеспечить безопасное развитие любых технологий в этой области срочно необходимо регулирование, предусматривающее запрет конкретных приложений или их использования.

С возникновением онлайн-пространств, управляемых искусственным интеллектом, существенно изменилась степень осознанного присутствия и активности человека в публичном пространстве. Свобода мысли и свобода совести неотделимы от способности человека к постоянной

95 Eileen Guo, [An AI chatbot told a user how to kill himself — but the company doesn't want to “censor” it](#) [Чат-бот с ИИ рассказал пользователю, как покончить с собой, но компания не хочет «подвергать его цензуре»] *MIT Technology Review*, 6 February 2025; Kate Payne, [In lawsuit over teen's death, judge rejects arguments that AI chatbots have free speech rights](#) [В судебном иске по делу о смерти подростка судья отклонил аргументы о том, что чат-боты с ИИ обладают правом на свободу слова], *Associated Press*, 21 May 2025.

96 Zara Abrams, [Using generic AI chatbots for mental health support: A dangerous trend](#) [Опасная тенденция: использование универсальных чат-ботов для психологической помощи]; American Psychological Association, 12 March 2025; [APA calls for guardrails, education, to protect adolescent AI users](#) [Американская психологическая ассоциация призывает к мерам по защите и обучению подростков], American Psychological Association, Press Release, 3 June 2025; Nathalie Koubayová, [Meet ChatPal, the European bot against loneliness](#) [ChatPal — европейский бот, который избавит вас от одиночества], *Algorithm Watch*, 22 May 2023.

эволюции и изменениям⁹⁷. Вместе с тем сегодня присутствие человека в онлайн-пространстве (например, в формате публикаций в Интернете) стало постоянным. Речь идет в том числе об общедоступных данных и оцифровке архивов газет. Например, опубликованное когда-то мнение сегодня может быть использовано против человека. Можно утверждать, что это ограничивает возможность и свободу человека развиваться — как в своих собственных глазах, так и в глазах других людей. Таким образом, человек может быть изначально лишен «права развиваться как личность», которое Шарль Малик, один из авторов Всеобщей декларации прав человека, назвал сущностью свободы.

Судебная практика начинает признавать «право на забвение» в контексте защиты права на неприкосновенность частной жизни. Оно позволяет человеку потребовать удалить информацию о себе, находящуюся в открытом доступе, кроме тех случаев, когда сохранение доступа к такой информации обусловлено общественными интересами⁹⁸.

Присутствие человека в онлайн-пространстве превратилось в так называемый цифровой след. Это не только то, что человек сознательно публикует онлайн, но и данные обо всех его действиях в Интернете, включая посещение сайтов, которые регистрируются с помощью файлов cookie. Эти данные позволяют создавать более подробный портрет пользователя, что затрагивает его автономию. При этом сами пользователи не осознают масштаб и последствия такого профилирования. Помимо персонализированного контента, оно может использоваться более широко, в частности, в системах наблюдения, а также создавать угрозу кибербезопасности или использоваться для преследования правозащитников, диссидентов или инакомыслящих в авторитарных режимах⁹⁹.

97 См., например, Hannah Arendt, *The Human Condition* [Состояние человека], University of Chicago Press: 1953.

98 «Выбор мер с учетом конкретных обстоятельств каждого дела зависит от таких факторов, как достоверность или неточность информации, ее значение для общественной дискуссии, ее историческое, исследовательское или статистическое значение, негативные последствия для частной жизни человека в связи с постоянной доступностью этой информации в Интернете, а также время, прошедшее с событий, упомянутых в статье, или с момента публикации информации», [Joint Factsheet: The Right to be Forgotten. ECtHR and CJEU Case-Law](#) [Совместный информационный бюллетень: право на забвение. Судебная практика ЕСПЧ и Суда ЕС], European Court of Human Rights and the EU Agency for Fundamental Rights, 30 October 2024.

99 См. [Digital Protection](#) [Защита в цифровом пространстве] на сайте организации Front Line Defenders.

5. Нейротехнологии и свобода мысли

Еще одна область, имеющая непосредственное отношение к свободе мысли, особенно в контексте ИИ, — это нейротехнологии. В этом разделе представлен обзор основных вызовов для свободы мысли на фоне развития этой области.

В основополагающем докладе Консультативного комитета Совета ООН по правам человека 2024 года¹⁰⁰ под понятием «нейротехнология» понимается «целый ряд устройств и систем, взаимодействующих с центральной нервной системой с помощью электрических, магнитных, оптогенетических и других средств». Некоторые из этих устройств и систем служат в первую очередь для понимания работы мозга, другие же представляют собой методы прямого вмешательства в психические процессы, направленного на восстановление утраченных функций и повышение когнитивных способностей.

Последние открытия в области нейротехнологий обеспечили более глубокое понимание работы мозга и способствовали разработке новых методов лечения неврологических или психических заболеваний. Это стало возможным благодаря масштабным исследованиям при поддержке государства¹⁰¹. Вместе с тем стремительное развитие нейротехнологий, в том числе с опорой на искусственный интеллект, служит основанием для серьезных опасений по поводу ментальной неприкосновенности, целостности и автономии личности.

100 Доклад Консультативного комитета Совета по правам человека «Воздействие, возможности и проблемы применения нейротехнологии в отношении поощрения и защиты всех прав человека», [A/HRC/57/61](#), 8 августа 2024 г., п. 4. Еще одно определение: «Нейротехнологии охватывают устройства и процедуры, используемые для доступа, мониторинга, исследования, оценки, оказания влияния и/или имитации структуры и функций систем нейронных сетей животных или человека». [Предварительное исследование технических и правовых аспектов целесообразности принятия нормативного акта об этических аспектах нейротехнологий](#), ЮНЕСКО, 6 апреля 2023 г.

101 Например, [Brain Research through Advancing Innovative Technologies \(BRAIN\) Initiative](#) [Проект по исследованию мозга на основе развития инновационных технологий] в США; [Human Brain Project](#) [Проект по изучению человеческого мозга] в ЕС; [China Brain project](#) [Проект по изучению мозга] в Китае. См. также Hermann Garden, David E Winickoff, Nina Maria Frahm, Sebastian Pfoth, Sebastian Pfoth, [Responsible innovation in neurotechnology enterprises](#) [Ответственные инновации в компаниях в области нейротехнологий], OECD Science, Technology and Industry Working Papers 2019/05, p. 15; или [International Brain Initiative](#) [Международная инициатива по изучению мозга].

В своей рекомендации 2019 года Совет ОЭСР признал, что «некоторые применения нейротехнологий ставят этические, правовые и общественные вопросы, учитывая понимание центральной роли мозга и когнитивных функций в концепциях идентичности, свободы мысли, автономии, неприкосновенности частной жизни и благополучия человека», а также заявил о «необходимости широкого общественного обсуждения наилучшего будущего нейротехнологий в обществе»¹⁰². В проекте Рекомендации по этическим аспектам нейротехнологий ЮНЕСКО¹⁰³ также рассматриваются этические, правовые и общественные вызовы и вопросы, связанные с применением нейротехнологий, в контексте прав и достоинства человека, включая «автономию, неприкосновенность частной жизни, психическую и физическую неприкосновенность, личную идентичность, свободу мысли, риск дискриминации, неравенство и проблемы демократии <...>». Кроме того, признается, что особая осторожность в отношении вмешательств в работу чрезвычайно сложной нервной системы обусловлена ее ролью в координации поведения и психических процессов, поскольку «она обеспечивает автономию личности, ее способность к самоопределению как морального субъекта, ответственность за свои действия, сотрудничество с другими, принятие во внимание коллективных решений и развитие».

Нейротехнологии уже давно применяются в медицине, при этом их взаимосвязь с инженерными и вычислительными науками, а также разработка медицинских и коммерческих приложений значительно усилили их влияние. Однако они не только способствуют развитию медицины, но и несут риски для свободы мысли и других прав человека¹⁰⁴.

«... беспрецедентные возможности, которые открывают [нейротехнологии] для вмешательства внешних субъектов в осуществление человеком своих прав, порождают огромные этические проблемы и ставят под сомнение само понимание основополагающих принципов прав человека». Нейротехнологии

¹⁰² Recommendation of the Council on Responsible Innovation in Neurotechnology [Рекомендация Совета по ответственным инновациям в области нейротехнологий], Organisation for Economic Cooperation and Development, OECD/LEGAL/0457, adopted on 11 December 2019.

¹⁰³ Проект Рекомендации об этических аспектах нейротехнологий, ЮНЕСКО, 9 апреля 2025 г.

¹⁰⁴ Доклад Консультативного комитета Совета по правам человека, A/HRC/57/61.

обладают социально разрушительным потенциалом, поскольку они, как правило: «а) позволяют раскрывать когнитивные процессы; б) позволяют прямо изменять психические процессы и мысли человека; с) не поддаются сознательному контролю со стороны человека или осознанию им; d) делают возможным внешний доступ к мыслям, эмоциям и психическим состояниям без предварительного согласия; е) подпитываются «нейроданными», которые необходимы для их собственного функционирования, корректировки и оптимизации; и f) собирают, анализируют и обрабатывают большие массивы персональных данных, имеющих крайне конфиденциальный характер».

— Доклад Консультативного комитета Совета по правам человека «Воздействие, возможности и проблемы применения нейротехнологии в отношении поощрения и защиты всех прав человека», [A/HRC/57/61](#), 8 августа 2024 г.

Помимо медицины, использование продуктов на основе нейротехнологии в приложениях и устройствах, играх, сервисах для мониторинга здоровья, медитации и даже интеграция нейротехнологий в наушники – все это позволило технологическим компаниям получить данные о мозге человека, что имеет последствия для свободы мысли¹⁰⁵. Как утверждает ЮНЕСКО, «компании могут использовать нейронные данные, полученные с помощью неинвазивных нейротехнологических устройств, в маркетинговых целях. Обнаруживая сигналы, связанные с нашими предпочтениями и антипатиями, эти компании могут влиять на поведение клиентов для получения максимальной прибыли. Это поднимает тревожные вопросы о системах слежения, маркетинговых тактиках и политическом влиянии на наши частные мысли и эмоции, что в конечном итоге угрожает демократии и основам общества»¹⁰⁶. Другие виды сенсорных технологий способны косвенно собирать данные о нашей нейронной активности: отслеживать взгляд, распознавать и анализировать голос, определять эмоции по выражению лица. В связи с их использованием для оценки психического состояния человека возникают вопросы¹⁰⁷.

105 Farahany, *The Battle for Your Brain* [Борьба за ваш мозг].

106 *Этика нейротехнологий*, сайт ЮНЕСКО.

107 *Проект Рекомендации об этических аспектах нейротехнологий*, ЮНЕСКО.

«Речь идет не просто о здоровье, а скорее о том, как мы рассматриваем личность, наше достоинство и способность полностью осуществлять свои права в контексте противостояния между потребностями в области здравоохранения и рыночными целями. С одной стороны, у нас есть серьезные задачи в здравоохранении, поскольку на заболевания нервной системы, неврологические заболевания и психические расстройства приходится треть наших расходов на здравоохранение <...> С другой стороны, на потребительском рынке становятся востребованными «нейронные данные» (или «данные мозга») за пределами медицинского сектора для цифрового фенотипирования, аффективных вычислений, нейрогейминга и нейромаркетинга. С таким использованием нейротехнологий связаны вопросы доверия общества, защиты ментальной неприкосновенности, стремительного технологического и экономического развития. Кроме того, следует отметить тот факт, что такие разработки практически не контролируются или контролируются ненадлежащим образом».

Hervé Chneiweiss, Ethics issues and global governance of neurotechnologies, in *The risks and challenges of neurotechnologies for human rights* [Этические вопросы и глобальное управление нейротехнологиями в сборнике «Нейротехнологии: риски и вызовы для прав человека»], UNESCO, University of Milan-Bicocca, and State University of New York (SUNY) Downstate, 2023, pp. 48-49.

В вышеупомянутом докладе Консультативного комитета Совета по правам человека делается вывод о том, что нейротехнологии уникальным образом влияют на права человека. В нем рассматриваются и объясняются основные риски для прав человека, включая право на свободу мысли, которые связаны с нейротехнологиями и их (потенциальным) использованием, включая следующие:

- Технологии, которые первоначально предназначались для помощи людям с неврологическими заболеваниями, сегодня разрабатываются и продаются как приложения для повышения когнитивных способностей и других немедицинских целей. Такие технологии позволяют пользователям управлять определенными элементами внешней среды с помощью мыслей или общаться с другими людьми, имеющими аналогичный имплантат. Несмотря на

то что такие медицинские решения могут быть инновационными, отсутствие надлежащих этических и правозащитных рамок для немедицинских устройств вызывает серьезные опасения, в частности, в отношении свободы мысли, отсутствия дискриминации, а также обеспечения равноправного доступа к таким решениям в медицинских целях. Кроме того, вызывает беспокойство соблюдение прав детей: например, в случае если в играх используются интерфейсы «мозг – компьютер», краткосрочные или долгосрочные последствия использования которых для (психического) здоровья недостаточно изучены, или в случае когда нейротехнологические продукты используются с целью улучшения интеллектуальных способностей детей.

- Несмотря на то что «чтение мыслей» пока невозможно, нейротехнологии с элементами ИИ способны делать все более подробные выводы о мыслях и психических состояниях человека, что позволяет создавать индивидуальные психологические профили. Специфическая озабоченность по поводу защиты неприкосновенности частной жизни возникает в связи с тем, что нейротехнологии способны делать детальные умозаключения об идентичности человека, в частности, личностных характеристиках, когнитивных способностях или сексуальной ориентации. Дополнительные риски связаны с использованием таких данных в системах правосудия или национальной безопасности в ходе допросов подозреваемых или свидетелей, в том числе с помощью детекторов лжи, использующих нейротехнологии. Таким образом, нарушается право не свидетельствовать против себя и создается риск наказания людей за их мысли. Это касается также ситуаций, когда технологии еще, возможно, не достигли такого уровня точности, чтобы правильно оценивать мысли человека. Раскрытие невысказанных мыслей без согласия человека противоречит принципам свободы мысли¹⁰⁸.
- Устройства для диагностики психических состояний уже применяются на рабочих местах, в частности, в экстремальных условиях, для обеспечения бдительности и предотвращения аварий

108 Доклад Консультативного комитета Совета по правам человека, [A/HRC/57/61](#).

из-за усталости¹⁰⁹. Хотя ставки здесь высоки, использование нейротехнологий вызывает серьезное беспокойство в отношении прав человека, в частности, прав работников и права на свободу мысли. Это может привести к использованию «нейронаблюдения» в целях повышения производительности¹¹⁰. Закон ЕС об ИИ предусматривает меры по решению этой проблемы в конкретных ситуациях¹¹¹. Более того, потребительское применение этих технологий, особенно в отсутствие четкого этического и правозащитного надзора и учитывая тот факт, что долгосрочные последствия их использования неизвестны, представляет риски для неприкосновенности личности и (психического) здоровья.

- В отсутствие надлежащего регулирования разработки и использования нейротехнологий особому риску подвергаются уязвимые группы, включая пожилых людей, детей, людей с инвалидностью и лиц, находящихся в заключении или в других условиях содержания под стражей. В частности, возникают вопросы по поводу согласия, которое всегда должно быть предварительным, добровольным, информированным, реальным, прозрачным, действительным и никогда не должно быть подразумеваемым. Например, учитывая медицинские задачи, для решения которых разрабатываются многие технологии, людей с инвалидностью и их представителей следует привлекать к процессам разработки, а их потребности, мнения и права как конечных пользователей должны занимать в этих процессах центральное место. Кроме того, обеспечение прав людей с инвалидностью означает предоставление им полного доступа к нейротехнологическим устройствам по завершении их разработки, а также проверки безопасности, эффективности и гарантий

¹⁰⁹ К ним относятся такие отрасли, как добыча полезных ископаемых, строительство, грузоперевозки, авиация, железные дороги. Farahany, *The Battle for Your Brain* [Борьба за ваш мозг]. См. также José M. Muñoz, Laura Isaza, and Tarini Mehta, [Tech is coming for your brain data: how a Chilean politician turbocharged the “neurorights” movement](#) [Технологическим компаниям нужны данные о вашем мозге: как чилийский политик запустил движение за «нейроправа»], *The Boston Globe*, 10 September 2024.

¹¹⁰ Доклад Консультативного комитета Совета по правам человека, [A/HRC/57/61](#).

¹¹¹ [Краткое содержание закона об ИИ](#), сайт ЕС, посвященный закону об ИИ, 27 февраля 2024 г., обновление от 30 мая 2024 г. (на англ. языке). В соответствии с законом распознавание эмоций на рабочих местах или в образовательных учреждениях, за исключением случаев, когда это необходимо в медицинских целях или в целях безопасности, относится к категории запрещенных систем ИИ. См. также Nora Santalu, [Neurotechnologies under the EU AI Act: Where law meets science](#) [Нейротехнологии в рамках закона ЕС об ИИ: где право соприкасается с наукой], IAPP, 12 May, 2025.

соблюдения прав человека на этапах проектирования, разработки и применения¹¹². Такой подход соответствует социальной модели, закрепленной в Конвенции о правах инвалидов, в которой, в отличие от дискриминационной медицинской модели, традиционно ориентированной на профилактику и лечение, люди с инвалидностью рассматриваются как часть человеческого разнообразия и участники жизни общества¹¹³.

Регулирование должно обеспечить защиту прав человека путем определения четких рамок на основе прав человека в области коммерческого использования нейротехнологий, при этом приоритетными должны оставаться достижения в сфере медицины.

112 Доклад Консультативного комитета Совета по правам человека, [A/HRC/57/61](#).

113 Международный альянс по проблемам инвалидности. Представление материалов по запросу Консультативного комитета Совета по правам человека на тему нейротехнологий и прав человека, 2 июля 2023 г. (на англ. языке), см. OHCHR [Neurotechnology and human rights webpage](#) [Нейротехнологии и права человека на сайте УВКПЧ].

6. Нормативно-правовая база

Регулирование ИИ на протяжении всего его жизненного цикла отстает от стремительных темпов развития одной из самых сложных и значимых областей в истории человечества. На региональном и мировом уровне существует ряд документов, не имеющих обязательной силы, но при этом нет ни одной всеобщей юридически обязательной нормы, обеспечивающей защиту прав человека¹¹⁴.

В принятой в 2021 году Рекомендации об этических аспектах ИИ Генеральная конференция ЮНЕСКО признала, что имеет место «существенное и динамичное позитивное и негативное воздействие, оказываемое искусственным интеллектом (ИИ) на общественные системы, окружающую среду, экосистемы и жизнь человека, в том числе его интеллектуальную деятельность, отчасти обусловленное появлением новых возможностей, позволяющих путем использования ИИ влиять на процессы мышления, взаимодействия и принятия человеком решений, а также на сферу образования, гуманитарные, социальные и естественные науки, культуру и информационно-коммуникационные технологии». Рекомендация в первую очередь направлена на защиту прав человека и человеческого достоинства на основе принципов прозрачности, справедливости и подконтрольности. Кроме того, в преамбуле документа содержится предупреждение о дисбалансах, связанных с ИИ, которые могут привести к серьезным рискам и потенциальным последствиям в разных сферах, а также призыв к укреплению взаимодействия и солидарного подхода в глобальном масштабе, включая многостороннее сотрудничество.

— [Рекомендация об этических аспектах искусственного интеллекта, ЮНЕСКО, 23 ноября 2021 года](#)

¹¹⁴ Искусственный интеллект, сайт ЮНЕСКО; проект Рекомендации об этических аспектах нейротехнологий; ЮНЕСКО, *Artificial intelligence* [Искусственный интеллект], сайт ОЭСР; Управление искусственным интеллектом в интересах человечества, Консультативный орган ООН по вопросам искусственного интеллекта, сентябрь 2024 г.; [Глобальный цифровой договор ООН](#).

Закон Европейского союза об искусственном интеллекте позволил добиться значительного прогресса, в частности, в области защиты прав человека. Поскольку закон был принят недавно и требует дополнительных разъяснений и развития инструментария, а также создания национальных механизмов по его имплементации, он еще не продемонстрировал свои возможности по снижению негативных последствий ИИ для прав человека. В 2024 году Совет Европы также принял Рамочную конвенцию об искусственном интеллекте и правах человека, демократии и верховенстве права, к которой присоединились новые участники¹¹⁵.

Несмотря на имеющиеся вопросы по поводу объема и эффективности средств защиты в этих документах¹¹⁶, важно обеспечить их ратификацию, внедрение и, если необходимо, укрепление этих инструментов. Для этого потребуются ресурсы и готовность на уровне государств, а также — в случае Закона ЕС об ИИ — со стороны институтов Европейского союза. Несмотря на то что отраслевые стандарты не являются обязательными и не заменяют меры регулирования, они также способствуют созданию этических рамок¹¹⁷. Кроме того, БДИПЧ отмечает появление научных работ и инициатив по вопросу принятия глобальных нормативных рамок, которые становятся основой для дискуссии по этой теме¹¹⁸.

115 Конвенция также открыта для подписания принимавшими участие в ее подготовке государствами, которые не являются членами Совета Европы. Разработаны специальные процедуры, позволяющие присоединиться к конвенции государствам, не являющимся членами Совета Европы. По состоянию на январь 2025 года конвенцию подписали США и Израиль. Чтобы конвенция приобрела обязательную юридическую силу, она должна пройти процедуру ратификации.

116 См. [EU: Artificial Intelligence rulebook fails to stop proliferation of abusive technologies](#) [Свод правил ЕС в области ИИ не остановил распространение технологий, нарушающих права человека] Amnesty International, 13 March 2024; и [ENNHRI Calls on Council of Europe member States to ensure strong human rights protection in the draft Convention on AI, Human Rights, Democracy and Rule of Law](#) [Европейская сеть НПЗУ призывает государства-члены Совета Европы обеспечить надежную защиту прав человека в проекте Конвенции об ИИ, правах человека, демократии и верховенстве права] ENNHRI, 17 May 2024. Всесторонний анализ этого закона и других нормативных актов в этой области см. Sandra Wachter, [Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This means for the European Union, the United States, and Beyond](#) [Ограничения и пробелы в Законе ЕС об ИИ и директивах об ответственности в области ИИ: что это означает для Европейского союза, США и других стран], Yale Journal of Law and Technology, Vol. 26, Issue 3, pp. 671-718; обсуждение конкретных нейротехнологий, а также пробелов в законодательстве ЕС см. Christoph Bublitz, [Banning biometric mind reading: the case for criminalizing mind probing](#) [Запрет биометрического чтения мыслей: аргументы в пользу криминализации распознавания мыслей], Law, Innovation and Technology, Vol. 16, Issue 2, pp. 432-462.

117 См., например, [IEEE Frameworks](#), Institute of Electrical and Electronics Engineers (IEEE) Tech Ethics [Стандарты Института инженеров в области электротехники и электроники].

118 См., например, Alexander Kriebitz, Caitlin C Corrigan (eds.), [Promoting and Advancing Human Rights in Global AI Ecosystems: The Need for A Comprehensive Framework under International Law](#) [Поощрение

Закон ЕС об ИИ — это на сегодняшний день самая полноценная попытка регулирования разработки и использования искусственного интеллекта, намного опережающая другие нормативные документы. Он устанавливает следующую риск-ориентированную категоризацию систем ИИ: запрещенные практики ИИ, высокорисковые системы ИИ, а также категории ограниченного и минимального риска. Закон предусматривает регулирование ИИ-систем общего назначения, которые работают на основе базовых моделей, таких как генеративный ИИ, а также закрепляет ряд обязательств, в частности, в отношении прозрачности и подотчетности. Кроме того, закон выделяет мощные модели ИИ общего назначения, представляющие системные риски, устанавливает дополнительные требования и уточняет, что такие модели могут использоваться в качестве высокорисковых систем или могут быть интегрированы в такие системы. В этой связи устанавливаются определенные обязательства по взаимодействию с такими системами для обеспечения их соответствия требованиям закона. Следует отметить, что закон предъявляет меньше требований к моделям ИИ общего назначения, которые распространяются на основе бесплатных открытых лицензий¹¹⁹. Закон ЕС об ИИ, использующий подход, основанный на оценке рисков, можно использовать в качестве модели в других регионах.

К запрещенным относятся ИИ-системы, которые предусматривают использование подсознательных, манипулятивных или вводящих в заблуждение методов для искажения поведения и затруднения принятия обоснованных решений, и причиняют значительный вред; использование уязвимостей, касающихся возраста, инвалидности или социально-экономического положения, для искажения поведения с причинением значительного вреда; использование систем биометрической категоризации (за исключением использования правоохранительными органами); осуществление социального скоринга на основе социального поведения или личностных характеристик; оценку риска совершения человеком уголовного преступления

и продвижение прав человека в глобальных ИИ-экосистемах: потребность во всеобъемлющих рамках в международном праве], February 2025.

119 Краткое содержание закона об ИИ, веб-сайт ЕС, посвященный закону об ИИ, 27 февраля 2024 г., обновление от 30 мая 2024 г. (на англ. языке).

исключительно на основе профилирования или личностных характеристик, за исключением случаев, когда она используется в дополнение к оценке, проведенной людьми, на основе объективных проверяемых фактов, непосредственно связанных с преступной деятельностью; формирование баз данных для распознавания лиц путем массового сканирования изображений лиц в Интернете или с записей камер видеонаблюдения; распознавание эмоций на рабочих местах или в образовательных учреждениях, за исключением случаев, когда это необходимо в медицинских целях или по соображениям безопасности; удаленную биометрическую идентификацию в режиме реального времени в общественных местах для целей правоохранительных органов, за некоторыми исключениями, которые требуют, в частности, оценки воздействия такой идентификации на основные права человека.

Поставщики систем ИИ высокого риска должны выполнять особые требования, которые касаются управления рисками, качества данных, прозрачности, подотчетности и человеческого надзора. Перед развертыванием, за некоторыми исключениями и при определенных условиях, в отношении таких систем также проводится оценка воздействия на основные права человека. ИИ-системы, которые используются для профилирования, всегда относят к высокорисковым.

—Краткое содержание закона об ИИ, веб-страница сайта ЕС, посвященная закону об ИИ, 27 февраля 2024 г., последнее обновление от 30 мая 2024 г. (на англ. языке).

До принятия закона об искусственном интеллекте в ЕС был принят пакет нормативно-правовых актов о цифровых услугах, которые регулируют работу онлайн-платформ и посредников, а также устанавливают правила для крупнейших платформ и поисковых систем. Кроме того, эти законы направлены на устранение системных рисков в отношении незаконного контента, свободы выражения мнений, свободы и плюрализма СМИ, дискриминации, защиты прав потребителей и прав детей, общественной безопасности и избирательных процессов, гендерного насилия, а также психического и физического благополучия. Они также закрепляют определенный

уровень прозрачности и надзора со стороны властей, как и возможность использования рекомендательных систем, которые не основаны на профилировании пользователей¹²⁰.

Эти нормативно-правовые акты Европейского союза регулируют чрезвычайно сложную область, и их влияние еще не проявилось в полной мере. Необходимые инструменты и институциональные рамки, особенно на национальном уровне, на момент подготовки публикации находятся в процессе разработки¹²¹. Кроме того, нет гарантий, что такое регулирование обеспечит защиту прав человека в контексте развития ИИ¹²². Законодательство ЕС играет важную роль, но его действие не распространяется на все государства ОБСЕ и другие регионы мира, поэтому необходимо регулирование на глобальном уровне¹²³. Закон Евросоюза об ИИ и Рамочная конвенция Совета Европы предусматривают исключения в целях национальной безопасности. Поскольку ЕС является самым передовым регулятором ИИ в мире, такие исключения вызывают серьезные вопросы по поводу соблюдения прав человека, а также могут сыграть роль в новой глобальной гонке вооружений, где лидерство в развитии ИИ обеспечивает превосходство¹²⁴. В таких условиях крайне важно развивать международное сотрудничество и создавать единые стандарты, которые позволят устранить такие пробелы, предотвратить злоупотребления и в то же время обеспечить поддержку ответственных инноваций.

120 European Commission, [DSA: Very large online platforms and search engines](#) [Закон о цифровых услугах: крупнейшие онлайн-платформы и поисковые системы](последнее обновление веб-страницы от 12 февраля 2025 г.); Европейская комиссия, [Пакет законов о цифровых услугах](#) (последнее обновление веб-страницы от 12 февраля 2025 г.) (на англ. языке).

121 См. [Towards meaningful fundamental rights impact assessments under the DSA](#) [Проведение содержательных оценок воздействия на основные права в соответствии с законом о цифровых технологиях], European Center for Not-for-Profit Law and Access Now, 15 September 2023; Aninda Chakraborty, [LatticeFlow AI unveils EU compliance framework for Generative AI](#) [LatticeFlow AI представляет систему соответствия требованиям ЕС к генеративному ИИ], *Tech Monitor*, 17 October 2024; [Обзор всех национальных планов по реализации положений Закона об ИИ](#), веб-сайт ЕС, посвященный Закону об ИИ, 8 ноября 2024 г. (последнее обновление от 19 мая 2025 г.) (на англ. языке).

122 [Withdrawal of the AI Liability Directive Proposal Raises Concerns Over Justice for AI Victims](#) [Отзыв предложения о принятии директивы об ответственности за использование ИИ поставил вопрос о справедливости для жертв искусственного интеллекта], Center for Democracy and Technology, 12 February 2025.

123 Нормативные документы были приняты в ряде стран, не входящих в ОБСЕ. Эти документы не являются предметом анализа в данной публикации. ЕС остается ведущим регулятором в этой сфере.

124 Iliara Carroza, Nicholas Marsh and Gregory M. Reichberg, [Dual-Use AI Technology in China, the US and the EU: Strategic Implications for the Balance of Power](#) [Технологии ИИ двойного назначения в Китае, США и ЕС: стратегические последствия для баланса сил], PRIO Paper, (Oslo: PRIO, 2022).

Разработка продуктов для медицинской сферы регулируется определенными нормативными актами, при этом нейротехнологии, создаваемые в коммерческих целях, создают новые риски для прав и основных свобод человека. Такие технологии требуют более эффективных механизмов защиты. Следует отметить, что на национальном или региональном уровне, а также в рамках закона ЕС об искусственном интеллекте есть отдельные — как правило, немногочисленные — элементы регулирования нейротехнологий, но универсальных международных и юридически обязательных стандартов в этой сфере не существует¹²⁵. В настоящее время в рамках продолжения работы по этике в сфере ИИ в ЮНЕСКО обсуждается первый глобальный стандарт по этическим аспектам в нейротехнологиях, не имеющий обязательной силы¹²⁶. Помимо этого, необходимо всестороннее регулирование, учитывающее сложность вызовов, связанных с нейротехнологиями, как в рамках действующих механизмов защиты, которые могут включать новые риски для прав человека, так и с помощью новых инструментов, касающихся разработки и применения нейротехнологий.

В своем докладе 2024 года¹²⁷ Специальный докладчик ООН по вопросу о праве на неприкосновенность частной жизни представил предложение об обновлении принятой в 1990 году резолюции Генеральной Ассамблеи 45/95 «Руководящие принципы регламентации компьютеризованных картотек, содержащих данные личного характера». В предложенном тексте говорится, что обработка нейроданных «не должна использоваться для манипулирования свободой мысли и совести людей или ее изменения, чтобы сделать их зависимыми от третьей стороны или повлиять на их представления, безопасность или автономию, а также их естественную мозговую идентичность или нейрокогнитивную целостность». Такие данные «также не могут обрабатываться в целях, не связанных с укреплением здоровья, диагностикой, реабилитацией и облегчением течения заболеваний в контексте осуществления права

¹²⁵ См., например, [Рекомендация Совета по ответственным инновациям в области нейротехнологий](#) OECD/Legal/0457, принята 11 декабря 2019 г. (на англ. языке).

¹²⁶ [Этические аспекты нейротехнологий: ЮНЕСКО назначила международную группу экспертов для подготовки нового глобального стандарта](#), пресс-релиз ЮНЕСКО, 22 апреля 2024 г.

¹²⁷ [A/79/173](#): Доклад Специального докладчика по вопросу о праве на неприкосновенность частной жизни Аны Браян Нугререс Предложение об обновлении резолюции 45/95 Генеральной Ассамблеи от 14 декабря 1990 года, озаглавленной «Руководящие принципы регламентации компьютеризованных картотек, содержащих данные личного характера», 17 июля 2024 г.

на здоровье или с проведением научных исследований в областях биологии, психологии и медицины для уменьшения страданий или улучшения состояния здоровья». В апреле 2025 года Совет ООН по правам человека принял [Резолюцию 58/6](#), в которой признал, что «продолжение работы над некоторыми видами их применения может вызывать ряд этических, правовых и социальных вопросов и иметь последствия для человеческого достоинства и автономии, что делает необходимым обеспечение эффективного уважения, защиты и реализации прав человека в этом контексте». Совет поручил Консультативному комитету разработать набор руководящих принципов применения существующей правозащитной базы к концептуализации, проектированию, разработке, тестированию, использованию и внедрению нейротехнологий¹²⁸.

В последние годы люди стали значительно лучше понимать, как социальные сети используют их данные, несмотря на сохраняющиеся вызовы и отсутствие четкого понимания последствий. В то же время осознание возможных рисков для прав человека, связанных с предоставлением своих нейроданных, крайне ограничено, в том числе из-за отсутствия нормативно-правовой базы¹²⁹.

ИИ-революция так называется, поскольку ИИ кардинально изменил жизнь человека и продолжит менять ее в будущем. Вместе с тем эта революция произошла без какой-либо публичной дискуссии или вклада общества в принятие решений. Ситуацию, в которой потребители присутствуют на цифровом рынке и только отчасти представляют себе имеющиеся у них возможности выбора и имеют лишь ограниченную способность повлиять на них, нельзя приравнять к демократическому обсуждению. Аналогичным образом стремительное развитие нейротехнологий и особенно коммерциализация устройств, которые взаимодействуют с мозгом человека или с нейронными цепями, оказывают беспрецедентное влияние на права человека. Государства-участники ОБСЕ несут обязательства по восстановлению условий для публичной дискуссии по значимым для общества вопросам. К ним относятся «улучшение когнитивных способностей» здоровых людей или даже «расширение человеческих возможностей», которые

¹²⁸ [Руководящие принципы прав человека в области нейротехнологий](#), сайт Совета ООН по правам человека (на англ. языке).

¹²⁹ Farahany, *The Battle for Your Brain* [Борьба за ваш мозг].

могут иметь настолько серьезные последствия для человечества, что заслуживают всестороннего обсуждения, по итогам которого могут быть приняты обоснованные решения о регулировании технологий такого типа. Круг участников таких обсуждений и процессов принятия решений должен быть намного шире и не должен ограничиваться крохотной группой специалистов, участвующих в разработке таких технологий или немногочисленных исследователей, работающих в этой сфере, которые могут утверждать, что у них имеется обоснованное мнение по этим вопросам¹³⁰. Несмотря на то что это сложные вопросы, последствия применения таких технологий для общества должны быть четко сформулированы и доведены до сведения общественности, чтобы обеспечить подлинное демократическое обсуждение вопроса о том следует ли использовать ИИ или отказаться от таких технологий.

«Ажиотаж вокруг ИИ <...> не утихает, поэтому кажется бессмысленным, и даже невозможным, что-либо иное, кроме уверенного продвижения к неизбежному господству искусственного интеллекта в будущем. Но каким бы правдоподобным это не казалось, это всего лишь ощущение. Оно не соответствует действительности, по крайней мере, пока. На самом деле существует множество альтернатив нынешней версии ИИ, множество способов построения нового мира. Но как и ИИ, они не являются неизбежными. Чтобы они стали реальностью, нужно для начала ответить на один вопрос: хотим ли мы жить в таком мире?»

— Kate Brennan, Amba Kak and Dr. Sarah Myers West, *Artificial Power: 2025 Landscape Report* [Способности ИИ: обзор за 2025 год], AI Now Institute, 3 June 2025, pp. 9 and 12.

¹³⁰ См. также OECD, *Recommendation of the Council on Responsible Innovation in Neurotechnology*, Recommendation No. 5: Enable societal deliberation on neurotechnology [Рекомендация Совета по ответственным инновациям в нейротехнологиях, Рекомендация № 5: Содействовать общественному обсуждению нейротехнологий].

7. Выводы и рекомендации

Необходимость защиты свободы мысли приобретает все большее значение на фоне растущего воздействия ИИ-технологий на самосознание человека и общества. Существующая система разработки и использования таких технологий усугубляет дисбалансы прав и полномочий и представляет угрозу для демократии. Мысли и решения человека должны быть защищены от новейших видов манипуляций, незаметных, но эффективных, а невысказанные мысли должны оставаться нераскрытыми для других. В отсутствие нормативных рамок для защиты прав человека, распространение технологий на основе ИИ, похоже, размывает — если не стирает — границы демократической ответственности и подотчетности. На государствах лежит обязанность уважать, защищать и соблюдать права человека. Для этого необходимо обеспечить регулирование на национальном и международном уровне. Несмотря на то что государства пока не приняли соответствующие меры, у них есть время это сделать, в частности, в отношении разработки и применения нейротехнологий. Такое регулирование позволит предотвратить масштабные негативные последствия.

На основе неисчерпывающего перечня критериев, предложенного Специальным докладчиком ООН по вопросу о свободе религии или убеждений для определения того, что можно рассматривать как недопустимое манипулирование мышлением в конкретных случаях, Сьюзи Алегре и Аарон Шулл составили список факторов для правовой оценки незаконного манипулирования с помощью технологий. Они отмечают, что право на свободу мысли в рамках системы защиты прав человека под эгидой ООН «возникло в ответ на зверства, совершенные нацистской Германией, и как признание катастрофических рисков, которые идеологическая обработка населения несет для прав человека в целом. Эффективная защита права на свободу мысли имеет решающее значение для будущего демократии и прав человека во всем мире». Они также предлагают «включить новые правовые критерии

в замечание общего порядка Комитета по правам человека с учетом ряда дополнительных факторов:

- масштаб и сфера применения практики;
- осведомленность общественности или соответствующих лиц о тактике влияния;
- понимание общественностью или соответствующими лицами последствий применения определенной практики;
- способность соответствующей практики обходить рациональное мышление;
- интенсивность и продолжительность воздействия;
- целенаправленное воздействие на определенные когнитивные искажения или уязвимости (особенно с использованием «инсайдерской информации», полученной с помощью данных);
- любой дисбаланс сил;
- наличие практической возможности сказать «нет»».

— Susie Alegre and Aaron Schull, *Freedom of Thought: Reviving and Protecting a Forgotten Human Rights* [Свобода мысли: восстановление и защита забытых прав человека], Center for International Governance Innovation, 2024, p. 12.

В свете изложенных аргументов в целях создания благоприятной среды для свободы мысли государствам-участникам ОБСЕ рекомендуется:

- **Принять на национальном уровне политику, законы и нормативные акты на основе прав человека, а также создать институциональную инфраструктуру** для обеспечения уважения, защиты и реализации права на свободу мысли, а также всех прав и основных свобод человека в контексте растущей роли технологий на основе ИИ или использующих ИИ, включая нейротехнологии. Для того чтобы обеспечить уважение человеческого достоинства и демократических принципов при регулировании нейротехнологий необходимо руководствоваться этическими и правозащитными соображениями, что предполагает, в частности, отказ от бизнес-модели социальных сетей, а также применение подходов на основе

оценки рисков и принятие разработанных отраслью мер по защите прав человека на протяжении всего жизненного цикла ИИ.

- **Внедрить методологии оценки воздействия на права человека** в качестве стандартной практики, создать правовые и институциональные рамки, необходимые для обеспечения широкого использования таких методологий. Такие инструменты должны уделять приоритетное внимание уважению, защите и соблюдению прав человека на протяжении всего жизненного цикла ИИ, регулярно пересматриваться и адаптироваться к быстрым темпам развития ИИ и новым вызовам в области прав человека. **Разработать инструменты для включения свободы мысли в оценку воздействия на права человека** любых технологий на основе ИИ, рассматривать это право как в качестве отдельно существующего, так и в связи с другими правами человека, включая право на свободу выражения мнения и на неприкосновенность частной жизни.
- **Собирать данные на протяжении всего жизненного цикла ИИ, оценивать соответствие стандартам прав человека (включая свободу мысли), потенциальные передовые практики и способы получения пользы от ИИ, а также доступность и равные возможности.** Несмотря на стремительное внедрение новых технологий на основе ИИ во многих сферах, в мире нет не только соответствующего регулирования, но и эффективного мониторинга в этой области, необходимого для понимания воздействия соответствующих технологий на права человека, прогнозирования последствий их использования и выявления новых возможностей. Государства обязаны осуществлять мониторинг прав человека и обеспечивать их соблюдение. Особое внимание необходимо уделять наиболее уязвимым группам, включая людей с инвалидностью, детей, пожилых людей и лиц, находящихся в местах лишения свободы.
- **Развивать и поддерживать новые исследования и концептуальное разъяснение стандартов** в целях поощрения и защиты свободы мысли, особенно в контексте новых технологий на основе искусственного интеллекта, включая международное сотрудничество в системе ООН (например, на основе принятия резолюций Совета по правам человека), использование механизмов ООН по правам человека (запрос на представление замечаний

общего порядка / подготовку докладов), а также посредством иных аналогичных способов.

- **Работать с системами образования** и готовить их к отражению новых моделей общественной жизни; помогать учащимся понять, что в мире, который все больше зависит от искусственного интеллекта, можно сохранить свободу мысли, если действовать на основе критического мышления, как личность, обладающая автономией и свободой воли. Для этого недостаточно программ по повышению цифровой грамотности: новые государственные учебные программы должны включать информацию, навыки и модели поведения, необходимые для подготовки нынешнего и будущих поколений к изменениям, которые затронут все сферы жизни. Кроме того, следует предусмотреть программы обучения и информационные кампании для взрослого населения.
- **Организовать во всех образовательных и учебных учреждениях подготовку на тему этики и прав человека для тех, кто работает с ИИ или принимает решения о его использовании на протяжении всего жизненного цикла, чтобы гарантировать соблюдение прав человека или укрепление их защиты.** Обучение в процессе получения образования (например, в университете) и непрерывное обучение для тех, кто уже работает, призваны гарантировать, что специалисты, которые разрабатывают и внедряют ИИ и новые технологии или управляют ими (аналитики данных, инженеры и другие), прошли подготовку по вопросам этических и правозащитных аспектов использования технологий, в том числе на тему свободы мысли. Специалисты, работающие с ИИ на протяжении его жизненного цикла (включая тех, кто принимает решения о приобретении и внедрении таких технологий), должны знать, как проводить оценку воздействия на права человека и интегрировать меры по защите прав человека в технологии искусственного интеллекта на этапе проектирования.
- **Защищать независимые СМИ и укреплять их роль в защите общественных интересов, включая защиту свободы мысли, свободы мнений и их выражения,** в частности, путем выполнения рекомендаций Представителя ОБСЕ по вопросам свободы СМИ по этой значимой теме¹³¹.

¹³¹ Подробнее см. Представитель ОБСЕ по вопросам свободы СМИ. В фокусе - искусственный интеллект и свобода слова: [рекомендации по разработке политики](#);

- **Способствовать развитию подлинной общественной дискуссии о последствиях применения ИИ и нейротехнологий с точки зрения этики и прав человека, в частности, права на свободу мысли.** Эти дискуссии должны быть структурированными и междисциплинарными, с привлечением различных заинтересованных сторон, включая тех, чьи права могут быть затронуты, или пользователей таких технологий. Чтобы повысить осведомленность общества об этих последствиях и предоставить необходимую информацию для широкой дискуссии, необходимо привлекать общественно значимые СМИ. Кроме того, в ходе дискуссий должна быть обеспечена прозрачность и подотчетность, а их результаты должны найти отражение в политических решениях.
- **Работать совместно с независимыми институтами ОБСЕ, включая БДИПЧ и Представителя ОБСЕ по вопросам свободы СМИ, чтобы обеспечить внедрение ИИ-технологий и их регулирование на протяжении жизненного цикла в соответствии с принципами прав человека.**

— SAIFE Resource Hub, а также запланированная работа в рамках инициативы Media and Big Tech [СМИ и крупнейшие технологические компании].



ОБСЕ

БДИПЧ

