

ОБЗОРНЫЙ ДОКЛАД

**От узнаваемости к
уязвимости: когда
соцсети перестают
работать для
правозащитников**

Март 2026



СПЕЦИАЛЬНЫЙ
ДОКЛАДЧИК
ООН ПО
ПОЛОЖЕНИЮ
ПРАВЗАЩИТНИКОВ

С СОЛИДАРНОСТЬЮ И НАДЕЖДОЙ

ОГЛАВЛЕНИЕ

Введение	03
Методология	04
Международное право и международные стандарты	05
Некорректная модерация контента и сокращение онлайн-присутствия	06
Выполнение репрессивных требований властей	11
Онлайн-атаки и неадекватная реакция платформ	13
Неравенство пользователей и их контента	17
Сложности во взаимодействиях с социальными сетями	19
Заключение	23
Рекомендации	25

ВВЕДЕНИЕ

С момента своего создания социальные сети стали незаменимым инструментом для тех, кто занимается защитой прав человека. Их используют для того, чтобы собирать информацию, рассказывать об успехах и неудачах, обнародовать доклады и свидетельства, создавать и укреплять сообщества и сети — и давать возможность высказаться тем, чьи голоса не слышны.

Однако в ходе своей работы и консультаций с людьми, занимающимися защитой прав человека, Специальная докладчица ООН по вопросу о положении правозащитников («Спецдокладчица») пришла к выводу о том, что те же социальные сети в значительной степени препятствуют правозащитной деятельности. Своими как минимум несовершенными практиками модерации, алгоритмическим подавлением контента, неадекватным реагированием на онлайн-харассмент и выполнением репрессивных требований со стороны Государств социальные сети способствуют — намеренно или ненамеренно — затыканию ртов, лишению права голоса и созданию рисков для правозащитников [1].

В свете этих соображений Спецдокладчица публикует настоящий обзорный доклад, представляющий собой анализ воздействия, оказываемого социальными сетями на реализацию права защищать права человека. Содержащиеся в нем выводы подчеркивают срочную необходимость того, чтобы компании, которым принадлежат социальные сети, начали выполнять свои обязанности по защите прав человека, а Государства — соблюдать свои обязательства в сфере прав человека и способствовать тому, чтобы цифровые пространства оставались безопасной и благоприятной средой для тех, кто защищает права человека.

[1] Спецдокладчица определяет правозащитника или правозащитницу как лицо, которое индивидуально или совместно с другими действует в целях поощрения или защиты прав человека в соответствии с Декларацией ООН о правозащитниках.

МЕТОДОЛОГИЯ

Настоящий обзорный доклад основан на исследовании и консультациях [2], проведенных аппаратом Специальной докладчицы ООН по вопросу о положении правозащитников с организациями гражданского общества, которые работают с людьми, защищающими и поощряющими права человека во всех регионах мира.

Охват настоящего доклада был ограничен имевшимися временем, ресурсами и доступом к источникам; документ не претендует на исчерпание темы. Он лишь отражает опыт организаций, которые нашли возможность внести свой вклад в процесс консультаций. К нему следует относиться как к предварительному обзору ключевых тенденций и рисков — а не как к исчерпывающему анализу всех факторов, связанных с социальными сетями, которые влияют на реализацию права на защиту прав человека.

Специальная докладчица ООН по вопросу о положении правозащитников довела до сведения компаний [Meta Inc.](#) и [X Corp.](#) свои соображения относительно их политики и ее воздействия на реализацию права на защиту прав человека в двух письмах от 30 декабря 2025 года, проинформировав их о том, что эти письма и полученные на них ответы будут включены в настоящий обзорный доклад. На момент его публикации ответов получено не было.

[2] Консультации проводились по «правилам Чатем-хауса».

1. МЕЖДУНАРОДНОЕ ПРАВО И МЕЖДУНАРОДНЫЕ СТАНДАРТЫ

Принятая на основе консенсуса Генеральной Ассамблеей ООН в декабре 1998 года Декларация ООН о правозащитниках [3] провозгласила, что «каждый человек имеет право, индивидуально и совместно с другими, поощрять и стремиться защищать и осуществлять права человека и основные свободы на национальном и международном уровнях». Это положение подкрепляется частью 1 статьи 12, которая гласит, что «каждый человек имеет право, индивидуально и совместно с другими, участвовать в мирной деятельности, направленной против нарушений прав человека и основных свобод», в то время как в соответствии со статьей 6 каждый имеет право распространять информацию, мнения и знания обо всех правах человека, а также привлекать внимание общественности к этим вопросам.

Статья 6 Декларации о правозащитниках развивает норму, содержащуюся в статье 19 Международного пакта о гражданских и политических правах, которая защищает право на свободу мнений и их выражения, определяемую как свобода «искать, получать и распространять всякого рода информацию и идеи, независимо от государственных границ, устно, письменно или посредством печати или художественных форм выражения, или иными способами по своему выбору». Свобода выражения мнений может ограничиваться в соответствии с частью 3 статьи 19, которая требует, чтобы такие ограничения были ясными и точными (будучи «установлены законом»), а также необходимыми для достижения целей «уважения прав и репутации других лиц», «охраны государственной безопасности, общественного порядка» или «здоровья или нравственности населения».

Основными субъектами международно-правовых обязательств в сфере прав человека являются Государства, которые обязаны не допускать нарушений прав человека в пределах своей территории или юрисдикции третьими лицами, в том числе частными предприятиями. Однако согласно Руководящим принципам предпринимательской деятельности в аспекте прав человека ООН, сами предприятия обязаны: (a) избегать оказания или содействия оказанию неблагоприятного воздействия и (b) предотвращать или смягчать неблагоприятное воздействие, которое непосредственно связано с их деятельностью, продукцией или услугами вследствие их деловых отношений. Для выполнения своей обязанности по соблюдению прав человека предприятия должны иметь (a) программное обязательство выполнять свои обязанности по соблюдению прав

[3] A/RES/53/144, опубликована на различных языках: <https://www.ohchr.org/en/special-procedures/sr-human-rights-defenders/declaration-human-rights-defenders-different-languages>

человека; (b) процедуру обеспечения должной заботы (due diligence) о правах человека в целях выявления, предотвращения, смягчения последствий и представления отчетности о том, каким образом они устраняют оказываемое ими воздействие на права человека; (c) процедуры, позволяющие возместить ущерб в связи с любыми видами неблагоприятного воздействия на права человека, оказанного ими самими или при их содействии. Обязанность соблюдать права человека является международным деловым стандартом, обязательным для всех предприятий, и существует вне зависимости от способности или готовности Государств выполнять их обязательства в сфере защиты прав человека.

2. НЕКОРРЕКТНАЯ МОДЕРАЦИЯ Контента и сокращение онлайн-присутствия

Правозащитники во всех регионах сообщают о целом ряде проблем, связанных с некорректной модерацией контента в социальных сетях, что оказывает серьезное негативное воздействие на защиту прав человека. К ним относятся удаление незапрещенного контента, теньевые баны, произвольные блокировки аккаунтов, непоследовательное применение правил сообщества и общая непрозрачность внутренних процедур.

2.1 УДАЛЕНИЕ ЛЕГИТИМНОГО ПРАВозащитного контента

В своих сообщениях, полученных при подготовке настоящего доклада, несколько организаций рассказали о случаях удаления незапрещенного правозащитного контента якобы за нарушение политики платформы.

Одна организация сообщила о том, как принадлежащими Meta платформами удалялись публикации, касающиеся защиты прав человека в **Палестине**, со ссылками на присутствие в них «лиц и организаций, представляющих опасность», и «контента, содержащего насилие», без каких-либо понятных объяснений и возможностей обжаловать удаление. Среди таких публикаций были свидетельства о внесудебных казнях, произвольных задержаниях и мирных протестах; удаления в первую очередь затрагивали контент правозащитников и журналистов, некоторые из которых не получали никаких уведомлений или объяснений подобных актов цензуры. Кроме того, правозащитники указывают на непоследовательность: так, публикации с критикой возбуждающего ненависть или экстремистского контента

удалялись, в то время как сам такой возбуждающий ненависть или экстремистский контент, ставший предметом критики, оставался онлайн [4]. Одна из организаций, занимающихся защитой цифровых прав, также последовательно фиксировала цензурирование и удаление контента, касающегося нарушений прав человека в Палестине, социальными сетями, включая Facebook, X и Instagram [5]. Еще одна организация сообщила об удалении контента, документировавшего потери во время войны в **Сирии**, включая публикации, которые могли быть использованы для привлечения виновных к международно-правовой ответственности [6]. В качестве оснований для удаления такого контента соцсети обычно ссылались на оправдание «лиц и организаций, представляющих опасность», и наличие «контента, содержащего насилие», или «тяжелого для восприятия контента». Фотографии жертв, особенно детей, часто маркировались как «эксплуатация детей».

Аналогичные проблемы фиксируются в других регионах и на других платформах. В TikTok контент, связанный с нарушениями прав человека в **Синьцзяне** и **Тибете**, регулярно маркировался, удалялся или ограничивался в последние годы якобы за нарушение правил сообщества [7][8]. Белорусская правозащитная организация сообщила, что ее публикации о жертвах политических репрессий в **Беларуси** иногда маркировались как содержащие недопустимый или политический контент в Facebook и Instagram; обжалование приводило к разным результатам, а никакие объяснения не предоставлялись [9]. Журналисты в **Камбодже** лишались доступа к своим аккаунтам в Facebook после публикаций о земельных спорах, коррупции и вырубках лесов, а во время выборов 2023 года аккаунты нескольких независимых СМИ были заблокированы, что стало причиной крайне существенного ограничения доступа к независимой информации [10]. Региональная правозащитная организация сообщила, что одна из ее платных информационных кампаний о нарушениях прав человека в странах АСЕАН была маркирована как недопустимая или «чувствительная» и удалена с платформ социальных сетей в **Филиппинах** и **Камбодже** — без каких-либо понятных объяснений и несмотря на ее беспрепятственное проведение в других странах Азии [11].

[4] Сообщение (A), полученное Спецдокладчицей.

[5] Сообщение (B), полученное Спецдокладчицей.

[6] Сообщение (C), полученное Спецдокладчицей.

[7] *The Guardian*, “Revealed: how TikTok censors videos that do not please Beijing”, Сентябрь 2019 г. <https://www.theguardian.com/technology/2019/sep/25/revealed-how-tiktok-censors-videos-that-do-not-please-beijing>

[8] *Radio Free Asia*, “TikTok deletes videos related to Uyghur human rights violations”, Ноябрь 2024 г. <https://www.rfa.org/english/uyghur/2024/11/05/uyghur-tiktok-censors-abroad/>

[9] Сообщение (D), полученное Спецдокладчицей.

[10] Сообщение (E), полученное Спецдокладчицей.

[11] Сообщение (F), полученное Спецдокладчицей.

2.2 ТЕНЕВЫЕ БАНЫ И ОГРАНИЧЕНИЕ ПОКАЗОВ

Теневые баны, из-за которых публикации пользователя становятся невидимыми для других без какого-либо предупреждения, также повсеместно встречаются в практике правозащитников.

Несколько организаций отмечали, что аккаунты палестинских правозащитников и контент, касающийся нарушений прав человека в **Палестине**, в значительной степени подвергаются теневым банам на платформах Meta, причем эта практика стала более распространенной с начала осуществляемого Израилем геноцида [12], хотя наблюдалась и ранее [13]. Обычно это означает, что публикации этих правозащитников, касающиеся ли они нарушений прав человека на местах, информирования международной общественности или привлечения финансирования для помощи жертвам, видны меньшему числу пользователей, что делает их работу менее эффективной. Пользователи, подвергающиеся таким банам, не получают никаких уведомлений или объяснений. Подобные случаи бывает трудно установить или верифицировать, поскольку действия Meta, касающиеся применения алгоритмов, отличаются почти полным отсутствием прозрачности [14][15].

Организации, занимающиеся защитой свободы выражений мнений, обращают внимание на то, что теневые баны в непропорционально значительной степени затронули феминисток-правозащитниц в **Ливане**, которые сообщали об удалении своих публикаций, особенно на арабском языке, при том что посты, разжигающие ненависть в отношении женщин, оставались нетронутыми [16][17]. Эти же практики коснулись женщин-правозащитниц, работающих в сфере сексуальных и репродуктивных прав [18].

В случае, задокументированном одной из предоставивших информацию для настоящего доклада организацией, объединение журналистов-мигрантов в **Чили** сообщило, что публикации в Instagram о публичном мероприятии против расизма, не появлялись в лентах пользователей. Несмотря на неоднократные обращения в

[12] Amnesty International, "Israel/OPT: Post-ceasefire: Israel's genocide in the occupied Gaza strip continues", 27 ноября 2025 г., <https://www.amnesty.org/en/documents/mde15/0527/2025/en/>

[13] SMEX, "Israel's Fierce Attack on Palestinian Narrative and Content", Июнь 2022 г., <https://smex.org/israels-fierce-attack-on-palestinian-narrative-and-content/>

[14] Сообщение (А), полученное Спецдокладчицей.

[15] Сообщение (В), полученное Спецдокладчицей.

[16] Сообщение (В), полученное Спецдокладчицей.

[17] SMEX, "Research Report: Confronting Structural Silencing: Challenges and Resistance Among Digital Feminist Activists in Lebanon", Апрель 2025 г., <https://smex.org/research-report-confronting-structural-silencing-challenges-and-resistance-among-digital-feminist-activists-in-lebanon/>

[18] Сообщение (С), полученное Спецдокладчицей.

в службу поддержки, понятного объяснения относительно того, почему посты скрывались, так и не было получено, а полный доступ к ним был разблокирован только после продолжительных обсуждений. Отсутствие прозрачности оставило членов объединения в неведении относительно причин ограничений и лишило их желания продолжать публикации в последующем [19].

Теневые баны применялись также по отношению к правозащитникам, специализирующимся на защите прав людей ЛГБТК+. В 2024 году новая политика Meta [20] ограничила распространение «политического контента» в Instagram и Threads — без каких-либо объяснений, что означают эти новые правила и как они будут применяться. На практике это изменение резко снизило видимость достоверных информационных публикаций нескольких Instagram-аккаунтов на тему защиты прав людей ЛГБТК+ [21], включая аккаунты американских организаций Human Rights Campaign и GLAAD. Это стало новым эпизодом в долгой истории теневых банов, применяемых Meta в отношении ЛГБТК+-контента, который часто маркируется как контент откровенно сексуального характера [22].

2.3 НЕКОРРЕКТНАЯ АЛГОРИТМИЧЕСКАЯ И АВТОМАТИЗИРОВАННАЯ МОДЕРАЦИЯ

Многие из описанных выше примеров некорректной модерации связаны с использованием автоматизированных систем модерации, в том числе функционирующих на основе ИИ. По мнению одной из неправительственных организаций, подобные технологии часто недостаточно точны, чтобы на них можно было полностью полагаться для различения законного и незаконного контента, что приводит к удалению значительного объема незапрещенных публикаций — в том числе размещаемых правозащитниками [23].

Еще одна организация приводит в пример неверную классификацию автоматическими фильтрами незапрещенных публикаций со свидетельствами о применении репрессий во время выборов 2024 года в **Венесуэле** как «призывов к

[19] Сообщение (A), полученное Спецдокладчицей.

[20] Instagram, “Continuing our Approach to Political Content on Instagram and Threads”, Февраль 2024 г., <https://about.instagram.com/blog/announcements/continuing-our-approach-to-political-content-on-instagram-and-threads>

[21] Accountable Tech, “Meta’s Political Content Limit Causes Steep Drop in Reach for Accounts”, Июль 2024 г., <https://accountabletech.org/research/metas-political-content-limit-causes-steep-drop-in-reach-for-accounts/>

[22] Human Rights Campaign, “Meta’s New Policies: How They Endanger LGBTQ+ Communities and Our Tips for Staying Safe Online”, Январь 2025 г., <https://www.hrc.org/news/metas-new-policies-how-they-endanger-lgbtq-communities-and-our-tips-for-staying-safe-online>

[23] Сообщение (G), полученное Спецдокладчицей.

насилию», что приводило к их блокировкам [24]. Та же организация сообщила, что аккаунт организации, занимающейся защитой репродуктивных прав в **Бразилии**, был заблокирован навсегда после того, как на нем была размещена информация о мизопростоле — лекарстве для лечения язвы желудка, в течение длительного времени используемого бразильскими женщинами для вызова аборта [25].

2.4 БЛОКИРОВКИ АККАУНТОВ И МАССОВЫЕ ЖАЛОБЫ

Люди и организации, занимающиеся защитой прав человека, также сталкиваются с блокировками и ограничениями аккаунтов, спровоцированными массовыми скоординированными кампаниями по направлению жалоб, которые организуют «тролли», акторы, связанные с властями Государств, правые движения и группы, выступающие против тех или иных прав.

Одна из организаций в **Центральной Америке** рассказала об использовании ботов и скоординированных кампаний для направления фейковых жалоб, что приводит к удалению контента и блокировкам аккаунтов пользователей-правозащитников. Такие инциденты часто происходят без предупреждения и без объяснений, а процедуры обжалования непрозрачны или вовсе недоступны, в особенности для пользователей, не имеющих прямого доступа к каналам «доверенных партнеров» (Trusted Partner) [26]. В **Бангладеш** правозащитницы-феминистки, защитники прав квир-персон и активисты со светскими взглядами сталкивались с аналогичными проблемами после того, как скоординированные кампании по массовой подаче жалоб приводили к автоматической блокировке их аккаунтов в Facebook и Instagram без реальных возможностей добиться пересмотра таких решений [27]. Одна из организаций рассказала, что в **Иране** Instagram-аккаунты активисток, занимающихся защитой прав женщин, блокировались или подвергались теневым банам после массовой подачи жалоб, инициированной, как предполагается, иранскими властями [28]. В **Бельгии** принадлежащий молодой правозащитнице Instagram-аккаунт, посвященный просвещению в сфере борьбы с расизмом, блокировался с 2023 по 2025 год семь раз из-за массовых жалоб групп, выступающих против прав человека. В конце концов это вынудило ее закрыть аккаунт и прекратить публикации [29]. Одна из организаций в **Малайзии** рассказала, что аккаунты в Facebook и Instagram НПО по защите прав женщин и двух ее сотрудниц блокировались без каких-либо объяснений после кампании травли. Хотя аккаунты были впоследствии

[24] Сообщение (А), полученное Спецдокладчицей.

[25] Сообщение (А), полученное Спецдокладчицей.

[26] Сообщение (Н), полученное Спецдокладчицей.

[27] Сообщение (А), полученное Спецдокладчицей.

[28] Сообщение (В), полученное Спецдокладчицей.

[29] <https://www.instagram.com/p/DHvS29aomok/>

восстановлены, этот случай повлиял на работу организации, которая была вынуждена пересмотреть свою политику цифровых коммуникаций [30].

3. ВЫПОЛНЕНИЕ РЕПРЕССИВНЫХ ТРЕБОВАНИЙ ВЛАСТЕЙ

По мере того как люди, использующие социальные сети для поощрения и защиты прав человека, сталкиваются со вмешательством со стороны государств, которые стремятся к подавлению инакомыслия и контролю над нарративами [31], некоторые компании часто выполняют требования властей о цензуровании контента, противоречащие международно-правовым нормам в сфере прав человека или способствующие слежке за пользователями со стороны властей.

3.1 ВЫПОЛНЕНИЕ ТРЕБОВАНИЙ ОБ ОСУЩЕСТВЛЕНИИ ЦЕНЗУРЫ

В сообщениях, полученных Спецдокладчицей при подготовке настоящего доклада, несколько организаций уделили особое внимание случаям, когда власти Государств оказывали давление на социальные сети, требуя привести их практики модерации в соответствие с конкретными политическими взглядами или осуществлять прямую цензуру в отношении правозащитников. Такое давление может принимать формы замедления (то есть снижения пропускной способности платформы), запрета рекламы на платформе или юридических санкций в виде, например, штрафов.

В **Турции**, где социальные сети стали основным источником новостной информации и незаменимым инструментом независимой журналистики, власти за последние годы усилили контроль за ними, внося целую серию изменений в закон об интернете. Одна из организаций сообщает об очень высокой доле выполнения требований об удалении контента, направляемых властями Турции, всеми основными платформами: во второй половине 2024 года она составила 95% у TikTok, 85% у X и 79% у Instagram [32]. В марте 2025 года, когда был арестован лидер оппозиции Экрем Имамоглу, власти потребовали блокировки сотен аккаунтов в соцсетях, принадлежащих в том числе правозащитникам, журналистам, студенческим объединениям, организациям гражданского общества и инициативам по защите прав женщин. X выполнил многие из этих требований под угрозой замедления, тогда как согласно заявлениям Meta, компания не стала

[30] Сообщение (G), полученное Спецдокладчицей.

[31] См. [A/HRC/38/35](#), п. 13, [A/80/341](#), пп. 12 и 76.

[32] Сообщение (G), полученное Спецдокладчицей.

предпринимать никаких действий [33] и, как утверждается, выплатила существенный штраф [34]. Организации, предоставившие информацию при подготовке настоящего доклада, также сообщили, что Meta [35], TikTok [36], YouTube [37] и X [38] были принуждены к назначению своих представителей в стране под угрозой все более серьезных юридических последствий — несмотря на предостережения о том, что это даст в руки турецким властям более жесткий контроль над их платформами [39].

В Индии независимая медиаплатформа с 300 000 фолловеров *Bolta Hindustan*, специализирующаяся на социальной проблематике, была заблокирована на YouTube в апреле 2024 года после направления соответствующего требования Министерством информации и телерадиовещания. Юридический отдел Google поначалу вообще не предоставил медиаплатформе каких-либо объяснений блокировки [40]. Активисты в сфере защиты свободы слова осудили блокировку, и в конце концов аккаунт на YouTube был восстановлен [41]. Аналогичным образом Twitter (в настоящее время X) подвергался критике за недопущение на платформу голосов из Кашмира по требованию индийских властей. В 2019 году Комитет по защите журналистов сообщил, что Twitter блокировал для пользователей, находящихся в Джамму и Кашмире во время блокировок СМИ и отключений интернета, — после того, как получил уведомление от индийских властей [42]. В Пакистане стримы с протестов белуджей в Facebook и YouTube замедлялись или просто удалялись без каких-либо объяснений [43].

[33] Meta, “Case Studies - Content alleged to violate Turkish Law”, Март 2025 г.,

<https://transparency.meta.com/reports/content-restrictions/case-studies>.

[34] Associated Press, “Meta is fined by Turkey after refusing to restrict content on Facebook and Instagram”, Апрель 2025 г., <https://apnews.com/article/meta-facebook-instagram-turkey-erdogan-59f16571e884fd6f0f9597a72771131f>

[35] Facebook, “An Update on Facebook in Turkey”, январь 2021 г.,

<https://about.fb.com/news/2021/01/an-update-on-facebook-in-turkey/>

[36] TikTok, “An update on TikTok in Turkey”, январь 2021 г.,

<https://newsroom.tiktok.com/an-update-on-tiktok-in-turkey?lang=en>

[37] YouTube, “An update on YouTube in Turkey”, декабрь 2020 г.,

<https://blog.youtube/news-and-events/update-youtube-turkey/>

[38] *Balkan Insight*, “X Appoints Representative in Turkey, Bowing to Govt Pressure”, май 2024 г.,

<https://balkaninsight.com/2024/05/23/x-appoints-representative-in-turkey-bowing-to-govt-pressure/>

[39] Сообщение (G), полученное Спецдокладчицей.

[40] Сообщение (F), полученное Спецдокладчицей.

[41] *Medianama*, “YouTube Reinstates Blocked and Demonetised Channels”, май 2024 г., <https://www.medianama.com/2024/05/223-youtube-reinstates-blocked-demonetised-channels/>

[42] *Committee to Protect Journalists*, “India uses opaque legal process to suppress Kashmiri journalism, commentary on Twitter”, Октябрь 2018 г., <https://cpj.org/2019/10/india-opaque-legal-process-suppress-kashmir-twitter/>

[43] Сообщение (F), полученное Спецдокладчицей.

3.2 СОЗДАНИЕ ВОЗМОЖНОСТЕЙ ДЛЯ СЛЕЖКИ СО СТОРОНЫ ВЛАСТЕЙ

Также вызывает беспокойство отсутствие прозрачности и контроля над персональными данными, которые правозащитники сообщают социальным сетям. В нескольких случаях, описанных одной из организаций, принадлежащие правозащитникам пользовательские данные передавались, как утверждается, властям без их ведома и согласия. Так, в **Гватемале** Meta передала пользовательскую информацию правозащитника — причем касающуюся не только того аккаунта, который интересовал власти, но и всех его аккаунтов в соцсетях. Сам правозащитник не получал об этом официального уведомления и узнал об этом только когда его проинформировал об этом надежный источник, обладавший фактической информацией [44]. Это не может не вызывать серьезного беспокойства, особенно там, где известно о преследованиях властями инакомыслящих. Процедуры, в рамках которых платформы получают запросы от властей и отвечают на них, остаются непрозрачными. В результате люди, занимающиеся защитой прав человека, могут подвергаться слежке, запугиванию или другим негативным последствиям, не зная о том, что их личные данные были переданы или раскрыты, и не имея доступа к эффективным механизмам обжалования.

4. ОНЛАЙН-АТАКИ И НЕАДЕКВАТНАЯ РЕАКЦИЯ ПЛАТФОРМ

Социальные сети стали средой, где воспроизводятся типичные практики харассмента и травли, с которыми многие правозащитники и так сталкиваются в своей работе, — в широком спектре цифровых форм. Онлайн-харассмент стал как более распространенным, так и более организованным: правозащитники регулярно становятся мишенями кампаний травли, клеветы и дезинформации, разворачиваемых сразу в нескольких социальных сетях, часто с использованием автоматизированных «ботов» [45]. Эта тактика, иногда дополняемая инструментами

[44] Сообщение (A), полученное Спецдокладчицей.

[45] Резолюция, принятая Советом по правам человека 4 апреля 2025 года, <https://docs.un.org/ru/A/HRC/RES/58/23>; UN Women, “Tipping point: The chilling escalation of online violence against women in the public sphere”, 2025, <https://www.unwomen.org/en/digital-library/publications/2025/12/tipping-point-the-chilling-escalation-of-violence-against-women-in-the-public-sphere-in-the-age-of-ai>; Front Line Defenders, “Global Analysis 2024/2025”, https://www.frontlinedefenders.org/sites/default/files/1609_fld_ga24-5_output.pdf.

на основе искусственного интеллекта, становится все более распространенной. При этом платформы социальных сетей оказались неспособны предложить адекватные механизмы для противодействия подобным атакам и защиты от них пользователей.

Чтобы подчеркнуть масштаб этой проблемы, организация Global Witness отметила в своем недавнем докладе [46] что 92% опрошенных ей правозащитников в сфере земельных прав и защиты окружающей среды сталкивались с онлайн-харасментом со стороны самых разных акторов, в том числе с участием физических лиц, ботов и в рамках организованных кампаний. Согласно данным исследования, правозащитники во всем мире чаще всего сталкивались с травлей в Facebook, за которым следуют X (Twitter) и Instagram [47].

4.1 ОНЛАЙН-ХАРАССМЕНТ, СВЯЗАННЫЙ С ГЕНДЕРОМ И ИДЕНТИЧНОСТЬЮ

Люди, которые уже сталкиваются с притеснениями в той или иной форме, обычно подвергаются бóльшим рискам и более значительным негативным проявлениям при использовании социальных сетей для защиты и поощрения прав человека. Во всех регионах мира правозащитники и правозащитницы, особенно женщины и представители ЛГБТИК+, сталкиваются с кампаниями дезинформации и клеветы, визуальной диффамацией, дипфейками, шантажом, угрозами, харасментом, доксингом и имперсонацией, на которые платформы реагируют недостаточно эффективно или не реагируют вовсе. Цели таких атак состоят в делегитимизации, запугивании, дискредитации или высмеивании правозащитников.

Одна из организаций в **Центральной Америке** обнаружила, что онлайн-атаки составляют около четверти всех посягательств на женщин-правозащитниц, задокументированных ею с 2020 по июнь 2025 года, причем это число значительно возрастает из года в год. В **Сальвадоре** онлайн-атаки составляют более половины всех посягательств на женщин-правозащитниц, и 90% этих атак происходят в социальных сетях [48].

Другая организация рассказала о том, как **сирийская** правозащитница Хиба Эззидин Аль-Хаджи стала жертвой кампании диффамации в Facebook, направленной против нее, ее семьи и организации, в которой она работает, — из-за публикации, в которой она выступила против браков по принуждению. Она получила множество угроз убийством в своем личном аккаунте, а также на странице организации, с призывами

[46] Global Witness, “Toxic platforms, broken planet”, 16 июля 2025 г., <https://globalwitness.org/en/campaigns/digital-threats/toxic-platforms-broken-planet/>.

[47] Там же.

[48] Сообщение (Н), полученное Спецдокладчицей.

к физической расправе над ней [49][50]. Несколькими годами ранее она уже становилась жертвой клеветы и угроз в Telegram [51].

Одна из организаций в **Бангладеш** сообщила, что политический кризис 2024 года стал переломным моментом, когда онлайн-харассмент от отдельных инцидентов перешел на уровень организованных кампаний с резким ростом использования доксинга, визуальной диффамации, языка ненависти на основе мемов, массовых жалоб на аккаунты в Facebook и Instagram и распространением адресной дезинформации. Такие атаки были часто связаны с темами гендера и религии, будучи особенно часто направлены против феминисток, защитников светского образа жизни и прав ЛГБТИК+ [52].

Несколько правозащитных организаций рассказали о том, как рост использования ИИ для генерирования дипфейков и искажения изображений повысил уязвимость правозащитников — в особенности женщин — онлайн [53][54][55]. Один из примеров — случай в **Венесуэле**, где правозащитница Марта Лиа Грахалес стала мишенью кампании дезинформации с использованием сгенерированного ИИ видео, которое якобы подтверждало ложный нарратив о том, что ее правозащитная работа — прикрытие для антиправительственного заговора. Видео широко распространилось в социальных сетях и в WhatsApp благодаря усилиям «цифровых войск», специально занимающихся организацией и активизацией таких атак. После распространения подобного контента его становится сложно опровергать или удалять [56]. Применительно ко всем вышеупомянутым примерам сообщалось, что реакция платформ на подобные инциденты была запоздалой или неэффективной.

Во многих ответах подчеркивалась роль социальных сетей в широком распространении и усилении эффекта вредоносного контента посредством алгоритмов, которые отдают приоритет сенсационному, поляризующему и показывающему насилие контенту, тем самым повышая охваты злоумышленников [57][58][59]. Одна из организаций указала, что ориентированные на максимизацию прибыли алгоритмы, используемые в настоящее время основными социальными

[49] Сообщение (А), полученное Спецдокладчицей.

[50] См. [AL SYR 3/2025](#).

[51] См. [AL OTH 129/2023](#) — письмо в адрес Telegram, на которое обладатели мандатов по специальным процедурам не получили ответа.

[52] Сообщение (А), полученное Спецдокладчицей.

[53] Сообщение (Н), полученное Спецдокладчицей.

[54] Сообщение (J), полученное Спецдокладчицей.

[55] Сообщение (А), полученное Спецдокладчицей.

[56] Сообщение (А), полученное Спецдокладчицей.

[57] Сообщение (I), полученное Спецдокладчицей.

[58] Сообщение (Н), полученное Спецдокладчицей.

[59] Сообщение (С), полученное Спецдокладчицей.

сетями, непосредственно способствуют цифровому насилию в отношении правозащитниц в Мезоамерике, которое представляет собой часть цикла насилия, часто принимающего и физические формы [60]. По данным вышеупомянутого исследования Global Witness, 84% защитников окружающей среды в Африке, Латинской Америке и Азии, сообщивших о преследованиях в связи со своей деятельностью, считают, что онлайн-харассмент «непосредственно» или «в какой-то степени» способствует преследованиям в офлайне [61].

4.2 ОТСУТСТВИЕ ЭФФЕКТИВНЫХ МЕХАНИЗМОВ ПОДАЧИ ЖАЛОБ

Авторы всех полученных Спецдокладчицей сообщений обращают особое внимание на отсутствие в распоряжении правозащитников быстро реагирующих, прозрачных или действующих живых людей систем подачи жалоб. В исследовании Global Witness только 12% правозащитников, направивших жалобы в социальные сети в связи с харассментом, были удовлетворены полученным ответом [62]. По сведениям одной из организаций, направивших сообщения Спецдокладчице, такие платформы как TikTok и Telegram вообще не отвечают на обращения за помощью от специализированных организаций в центральноамериканском регионе. Кроме того, многие процедуры требуют свободного владения английским языком и задействуют боты или автоматические ответы [63]. Одна из организаций в **Бангладеш** сообщила, что более 90% жалоб на онлайн-харассмент правозащитников в Facebook либо игнорировались, либо приводили к негативным последствиям для авторов жалоб, а не для тех, на кого они жаловались, — вплоть до блокировок аккаунтов [64]. Одна из правозащитниц в Либерии рассказала о том, как отсутствие реакции социальных сетей на случаи онлайн-атак на правозащитниц, борющихся за участие женщин в политической деятельности, в некоторых случаях приводило к реальным нападениям на них [65]. Организация, специально выделяющая время и ресурсы на службу цифровой безопасности для правозащитников, сообщила, что по большинству жалоб, которые она отслеживают, в конечном счете принимается то или иное решение. Однако продолжительное время ожидания ответа (до 14 дней), а в некоторых случаях также отсутствие ответа или вынесение отрицательного решения по жалобам подвергают правозащитников еще большей опасности. Кроме того, важно отметить, что у большинства правозащитников нет доступа к такого рода сопровождению и ресурсам, в результате чего их взаимодействия с социальными сетями гораздо реже заканчиваются положительным разрешением проблемы.

[60] Сообщения (B), (F), (H) и (J), полученные Спецдокладчицей.

[61] Global Witness, “Toxic platforms, broken planet”, 16 июля 2025 г.

[62] Там же.

[63] Сообщение (H), полученное Спецдокладчицей.

[64] Сообщение (A), полученное Спецдокладчицей.

[65] Сообщение (K), полученное Спецдокладчицей.

Организация также выразила обеспокоенность случаями имперсонации аккаунтов (когда за их владельцев выдают себя другие лица) и харассмента в чатах — это весьма распространенные проблемы среди женщин-правозащитниц на **Ближнем Востоке и в Северной Африке**, которые могут приводить к весьма серьезным негативным последствиям. Тем не менее, как сообщает организация, социальные сети регулярно не принимают никаких мер в таких случаях, просто предлагая пользователю заполнить анкету, что не приводит ни к какому реальному результату [66].

5. НЕРАВЕНСТВО ПОЛЬЗОВАТЕЛЕЙ И ИХ КОНТЕНТА

Несколько организаций обратили внимание на дискриминационное отношение социальных сетей к правозащитникам и публикуемому ими контенту, в особенности в привязке к их языкам и регионам [67].

5.1 ПРОЯВЛЕНИЯ ЯЗЫКОВОГО НЕРАВЕНСТВА

Одна из важнейших проблем состоит в нехватке языковых ресурсов в модераторских командах и инструментах. Это особенно очевидно для модерации контента не на английском, французском или испанском языках, когда опасный и хейтерский контент часто остается незамеченным и не вызывает никакой реакции.

Проведенное одной из организаций исследование онлайн-харассмента в **Бангладеш** показало, что такие платформы как Facebook неспособны обнаруживать недопустимый контент на языке бенгали и реагировать на него. Несмотря на многочисленные жалобы, аккаунты с хейтерским контентом остаются активными, а контент с уничижительными терминами типа «шахбаги» [68], используемыми хейтерами в отношении правозащитников и активистов, не маркируется и не удаляется [69]. В то же время чрезмерно активная модерация контента на других

[66] Сообщение (C), полученное Спецдокладчицей.

[67] Сообщения (A), (B), (C), (F) и (I), полученные Спецдокладчицей.

[68] Термин возник во время протестов 2013 года в районе Шахбаг столицы Бангладеш Дакки; используется как уничижительный эпитет для людей левых, либеральных или светских убеждений, приписывающий им исламофобию или приверженность «кампаниям общественного порицания» или самосуду.

[69] Сообщение (A), полученное Спецдокладчицей.

языках создает иные, но не менее вредоносные тенденции. Несколько организаций отметили, что посты на арабском языке, особенно касающиеся **Палестины, Сирии** и проблем региона в целом, модерируются более агрессивно, чем публикации на других языках [70].

В целом контент не на английском, французском или испанском языках удаляется или маркируется непропорционально часто — в связи с тем, что социальным сетям не хватает языковых ресурсов для модерации [71]. Аналогичную проблему затронула правозащитная организация в **Камбодже**, отметив, что удаления публикаций на кхмерском не сопровождаются никакими объяснениями, а для правозащитников, владеющих исключительно кхмерским, не существует никакой службы поддержки [72]. Еще одна организация рассказала, что язык урду стал «слепой зоной» для платформ в том, что касается языка ненависти и угроз в отношении правозащитников в **Пакистане** [73]. В **Иране** модерация фарсиязычного контента в Instagram также сопровождается проблемами в применении политики платформы при модерации контента на не являющихся западными языках [74].

5.2 ПРОЯВЛЕНИЯ РЕГИОНАЛЬНОГО НЕРАВЕНСТВА

Несколько организаций также сообщили о серьезных проявлениях неравенства между регионами в том, как социальные сети относятся к правозащитникам и их контенту [75]. Согласно результатам исследования, проведенного одной из организаций, хотя частота обращений правозащитников в службы поддержки в связи с причинением им того или иного вреда более-менее одинакова в разных регионах мира, африканские правозащитники реже получают ответ на свои жалобы, чем их европейские коллеги (50% и 72% случаев соответственно) [76]. Аналогичные пробелы в реагировании на случаи харассмента в отношении правозащитников в разных регионах затронула и другая организация [77].

[70] Сообщения (A), (B) и (C), полученные Спецдокладчицей.

[71] Сообщения (A), (B), (C) и (E), полученные Спецдокладчицей.

[72] Сообщение (E), полученное Спецдокладчицей.

[73] Сообщение (B), полученное Спецдокладчицей.

[74] Article 19, “Iran: Meta must overhaul Persian-language content moderation on Instagram”, Июнь 2022 г. <https://www.article19.org/resources/iran-meta-persian-language-content-moderation-instagram/>

[75] Сообщения (A), (B), (F) и (I), полученные Спецдокладчицей.

[76] Сообщение (I), полученное Спецдокладчицей.

[77] Сообщение (A), полученное Спецдокладчицей.

6. СЛОЖНОСТИ ВО ВЗАИМОДЕЙСТВИЯХ С СОЦИАЛЬНЫМИ СЕТЯМИ

6.1 НЕБЛАГОПРИЯТНЫЕ ДЛЯ ПРАВООЩИТНИКОВ РЕЗУЛЬТАТЫ

Сталкиваясь с вышеописанными притеснениями, нарушениями прав и предубеждениями, сами правозащитники обладают весьма ограниченным доступом к механизмам разрешения возникающих конфликтов.

Исследование, проведенное одной из организаций, показало, что в отсутствие специализированной помощи со стороны структур гражданского общества, сопровождающих правозащитников при направлении ими жалоб на платформы социальных сетей эффективность реакции последних остается крайне ограниченной. В правилах сообщества таких платформ как Meta многие формы онлайн-харассмента часто не считаются противоречащими их политике, особенно когда вредоносный контент преподносится как часть общественного обсуждения или как реализация свободы слова. В результате многие жалобы на харассмент игнорируются или не признаются основательными, а правозащитники не имеют доступа к какому-либо реальному механизму защиты. Даже в тех случаях, когда платформы все же принимают какие-то меры, их реакция часто запаздывает и не соответствует серьезности атак [78].

Другая организация определила, что основными препятствиями для получения женщинами-правозащитницами защиты в случаях онлайн-атак в **Мезоамерике** являлись отсутствие адекватных механизмов, сложности с получением доступа к существующим механизмам, особенно для не владеющих английским языком, длительное ожидание ответа, отсутствие последующего контроля и тот факт, что ответы часто автоматизированы и направляются с использованием ботов, при этом добиться общения с живым человеком практически невозможно [79]. В результате некоторые правозащитники предпочитают вообще не сообщать о случаях харассмента, с которыми они сталкиваются. 65% правозащитниц, сообщивших этой организации об онлайн-атаках, не сообщали о них самим платформам — в основном по причинам неэффективности используемых теми механизмов [80]. Исследование, проведенное еще одной организацией в **Бангладеш**, демонстрирует похожие результаты: по словам опрошенных правозащитников, механизмы обжалования в Facebook часто оказываются запутанными, непрозрачными и демотивирующими, в

[78] Сообщение (А), полученное Спецдокладчицей.

[79] Сообщение (Н), полученное Спецдокладчицей.

[80] Сообщение (Н), полученное Спецдокладчицей.

результате чего многие теряют последние иллюзии и больше не рассчитывают на подачу официальных жалоб [81].

В некоторых случаях, когда речь шла о прямых угрозах или использовании изображений правозащитников без их разрешения, платформы иногда реагировали, удаляя посты или блокируя аккаунты нарушителей. Однако такие вмешательства происходят редко и непоследовательно. Так же редко удается добиться благоприятного исхода, когда речь идет о временных или постоянных блокировках.

Непрозрачность процедур обжалования, отсутствие транспарентности и недоступность общения с живыми людьми — вот основные факторы, из-за которых правозащитники редко получают ответы и защиту, в которой они нуждаются. В результате многие правозащитники обращаются за помощью в специализированные организации гражданского общества, чтобы попробовать добиться нужного результата.

6.2 ОРГАНИЗАЦИИ ГРАЖДАНСКОГО ОБЩЕСТВА КАК ПОСРЕДНИКИ: ПРОБЛЕМНАЯ И НЕНАДЕЖНАЯ МОДЕЛЬ

Некоторые располагающие достаточными ресурсами правозащитные организации устанавливают прямые каналы связи с платформами социальных сетей, чтобы сообщать им о важных случаях и оказывать помощь правозащитникам. Хотя такие каналы способствуют достижению относительно более удовлетворительных результатов и ускорению процесса, организации гражданского общества часто называют их ненадежными или непоследовательными [82]. Одна из организаций сообщила, что взаимодействие с социальными сетями от имени правозащитников часто оказывается непродуктивным и не позволяет добиться реальных системных изменений [83]. Еще одна организация аналогичным образом указала, что результат никогда не может быть гарантирован. Когда эта организация попыталась помочь правозащитнику в **Казахстане** восстановить свой аккаунт, напрямую связав его с Meta, взаимодействие так и не состоялось, несмотря на повторные попытки [84].

Даже такие инициативы как программа Meta «Доверенный партнер» (Trusted Partner Program), специально созданная для налаживания сотрудничества между платформой и организациями гражданского общества, приводят к непоследовательным результатам. Одна из организаций рассказала, что в случаях, когда опубликованный правозащитниками контент удалялся или маркировался, она

[81] Сообщение (A), полученное Спецдокладчицей.

[82] Сообщения (A), (B), (C), (F) и (I), полученные Спецдокладчицей.

[83] Сообщение (I), полученное Спецдокладчицей.

[84] Сообщение (B), полученное Спецдокладчицей.

пыталась наладить взаимодействие по этим официальным каналам. Хотя ответы по каналу «Доверенный партнер» приходили от Meta чаще, чем от X или Telegram, во многих случаях они представляли собой автоматически сгенерированные тексты, не содержащие объяснений по существу. В тех же случаях, когда какие-то объяснения приводились, они редко свидетельствовали о признании или понимании более общего контекста той или иной ситуации, и передача вопроса на рассмотрение более высокой инстанции не гарантировало благоприятного разрешения [85].

В целом взаимодействие между организациями гражданского общества и платформами отличается непоследовательностью, отсутствием прозрачности и растущее нежелание взаимодействовать со стороны компаний. Несмотря на это, НПО продолжают стремиться к устранению пробела, создаваемого платформами, и все больше берут на себя посредническую роль, документируя причиняемый вред, вступая во взаимодействие с платформами и отстаивая права пользователей [86]. Sin embargo, esto no es un modelo sostenible, ya que asigna la carga de la protección y el apoyo a la sociedad civil, mientras persiste la falta de mecanismos accesibles y efectivos para las personas defensoras de los derechos humanos que están en situaciones de riesgo [87].

6.3 ОСЛАБЛЕНИЕ ПРИВЕРЖЕННОСТИ ПЛАТФОРМ ЗАЩИТЕ ПРАВ ЧЕЛОВЕКА

В последние годы наблюдается процесс, который можно считать не чем иным, как признаком ослабления приверженности платформ защите прав человека: основные социальные сети постепенно отказываются от сотрудничества с правозащитниками и НПО.

Несмотря на продолжительный опыт непосредственного сотрудничества с социальными сетями — как неофициально, так и в качестве официального участника различных программ «Доверенный партнер», — одна из организаций указала, что это взаимодействие становится все менее и менее эффективным, а время ожидания ответов в 2025 году значительно увеличилось. Вредоносный контент остается неудаленным до 10 или 15 дней после направления жалобы, даже в случаях серьезных угроз в отношении женщин-правозащитниц, на **Ближнем Востоке** и в **Северной Африке**. Также сообщается, что социальные сети в настоящее время гораздо менее восприимчивы к общественному давлению, чем ранее [88]. Другие организации гражданского общества также признают, что приверженность платформ защите прав человека заметно снизилась за последние три года — как и

[85] Сообщение (A), полученное Спецдокладчицей.

[86] Сообщения (A), (B) и (C), полученные Спецдокладчицей.

[87] Сообщение (A), полученное Спецдокладчицей.

[88] Сообщение (C), полученное Спецдокладчицей.

их интерес ко взаимодействию с гражданским сектором [89].

Эти изменения происходили параллельно с массовыми увольнениями компаниями, владеющими социальными сетями, ключевого персонала в сферах обеспечения доверия, безопасности пользователей, прав человека и противодействия дезинформации, а также с упразднением консультативных органов и программ, занимавшихся этой проблематикой. Это происходило на платформах X [90], Meta [91], YouTube [92] и TikTok [93]. Такие структурные изменения резко сузили возможности для диалога, имевшиеся ранее у правозащитников и организаций гражданского общества, и оказали негативное влияние на способность платформ разрешать как индивидуальные конфликтные ситуации, так и структурные проблемы [94]. Одновременно большинство крупных социальных сетей отказалось от ранее принятых важнейших принципов («политик»), вносящих свой вклад в защиту прав пользователей. Meta, X и YouTube ослабили защиту приватности пользователей, чтобы использовать их данные для обучения своих ИИ-инструментов. Meta отказалась от независимого фактчекинга в пользу «модерации силами сообщества» или автоматизированной модерации, что вызывает беспокойство по части распространения дезинформации. X и YouTube также отказались от политик по предотвращению дезинформации в ходе избирательных кампаний [95].

Хотя Meta с 2022 года публикует ежегодные отчеты по соблюдению прав человека, рассказывая о том, как компания выявляет и смягчает риски в сфере прав человека на своих платформах, в последнем отчете [96] не содержится никакой содержательной информации о вышеописанных проблемах. Другие платформы — X, TikTok или YouTube — не публикуют отдельных специализированных отчетов о своей деятельности в сфере защиты прав человека.

[89] Сообщения (I) и (J), полученные Спецдокладчицей.

[90] *Amnesty International*, “Global: Twitter’s decision to disband safety council threatens wellbeing of users”, Декабрь 2022 г., <https://www.amnesty.org/en/latest/news/2022/12/global-twitters-decision-to-disband-safety-council-threatens-wellbeing-of-users/>; *Electronic Frontier Foundation*, “Monetization, Not Human Rights or Vulnerable Communities, Matter Most at Twitter Under Musk”, Ноябрь 2022 г., <https://www.eff.org/deeplinks/2022/11/twitters-monetizable-users-not-human-rights-matter-most-under-musks-rein-leaving>; Сообщение (F), полученное Спецдокладчицей.

[91] *Free Press*, “Big Tech Backslide: How Social-Media Rollbacks Endanger Democracy Ahead of the 2024 Elections”, Декабрь 2023 г., <https://www.freepress.net/big-tech-backslide-how-social-media-rollbacks-endanger-democracy-ahead-2024-elections>; Сообщение (F), полученное Спецдокладчицей.

[92] *Free Press*, там же.

[93] *Business & Human Rights Centre*, “Germany: TikTok content moderators strike over mass layoffs & plans to replace them with AI”, Август 2025 <https://www.business-humanrights.org/en/latest-news/deutschland-content-moderatorinnen-protestieren-gegen-stellenabbau-geplante-ersetzung-durch-ki-algorithmus/>

[94] Сообщение (C), полученное Спецдокладчицей.

[95] *Free Press*, там же.

[96] *Meta*, [2024 Human Rights Report](#), Декабрь 2025 г.

7. ЗАКЛЮЧЕНИЕ: ПОСЛЕДСТВИЯ ДЛЯ ДЕЯТЕЛЬНОСТИ ПРАВОЗАЩИТНИКОВ ОНЛАЙН И ИХ СТРАТЕГИЙ

Описанные выше формы цифровых репрессий приводят к серьезным последствиям в виде сокращения цифрового пространства для гражданского общества, когда правозащитники постепенно лишаются возможности быть услышанными онлайн.

Сталкиваясь с ограничениями по части узнаваемости и заметности, харассментом или слежкой в социальных сетях, а также с отсутствием эффективной или адекватной поддержки со стороны платформ, которая могла бы исправлять подобные ситуации, многие правозащитники говорят о том, что они испытывают чувство страха и обеспокоенности за самих себя и за свои сообщества, как это видно из сообщений, полученных при подготовке настоящего доклада, [97] и из бесед Спецдокладчицы с правозащитниками. В результате их выбор невелик: они могут либо смириться со своей незаметностью, либо избегать острых тем — либо просто уйти из социальных сетей.

Женщины-правозащитницы в **Сальвадоре** снизили свою публичную активность из-за постоянного онлайн-харассмента [98]. Исследование, проведенное организацией Media Matters for Democracy в 2023 году в **Пакистане**, показало, что восемь из десяти женщин-журналисток практиковали самоцензуру онлайн, опасаясь цифрового и физического насилия [99]. Еще одна организация установила, что в **Камбодже** треть журналистов избегают публикаций контента с критикой властей в социальных сетях [100]. Самоцензура по острым политическим темам распространена и в других странах **Азии** [101]. Почти половина опрошенных организацией Global Witness правозащитников в сфере земельных прав и защиты окружающей среды сообщили, что харассмент, которому они подвергались онлайн, снизил их работоспособность и мешал им заниматься активизмом по этим вопросам [102].

Говоря в целом, основными последствиями стали менее заметное присутствие

[97] Сообщения (A), (H), (I) и (J), полученные Спецдокладчицей.

[98] Сообщение (H), полученное Спецдокладчицей.

[99] Сообщение (J), полученное Спецдокладчицей.

[100] Сообщение (E), полученное Спецдокладчицей.

[101] Сообщение (F), полученное Спецдокладчицей.

[102] Global Witness, там же.

правозащитников онлайн и меньшая публичность их деятельности — особенно это коснулось тех, кто принадлежит к маргинализированным сообществам, и тех, чья работа связана с чувствительными темами. Это создает предпосылки для формирования онлайн-экосистемы, в которой существующие тенденции только усугубляются: сообщества, которые уже сейчас больше других уязвимы для атак, рискуют полным изгнанием из цифрового пространства.

Некоторые правозащитники и правозащитные организации изучают альтернативные, более автономные платформы, на которых безопасности и приватности отдается приоритет перед охватами. Другие видоизменяют свои стратегии в том, что касается онлайн-присутствия, ограничивая его, используя псевдонимы или кодовые сообщения, избегая публикаций в реальном времени и так далее [103]. Однако крупные социальные сети остаются важнейшим инструментом для большинства правозащитников. Более новые платформы пока сильно отстают в том, что касается охватов и активности пользователей, и доминирующее положение в цифровой экосистеме в настоящее время сохраняют несколько крупных платформ, цель которых состоит в максимизации прибыли. Это заставляет правозащитников постоянно искать компромисс между заметностью и безопасностью. Несколько организаций рассказали нам о том, насколько правозащитникам сложно полностью отказаться от присутствия в соцсетях, не подвергая риску результативность своей работы, — и как отсутствие безопасных альтернатив продолжает создавать для них серьезные риски [104].

[103] Сообщения (A), (C), (F) и (I), полученные Спецдокладчицей.

[104] Сообщения (A), (B) и (C), полученные Спецдокладчицей.

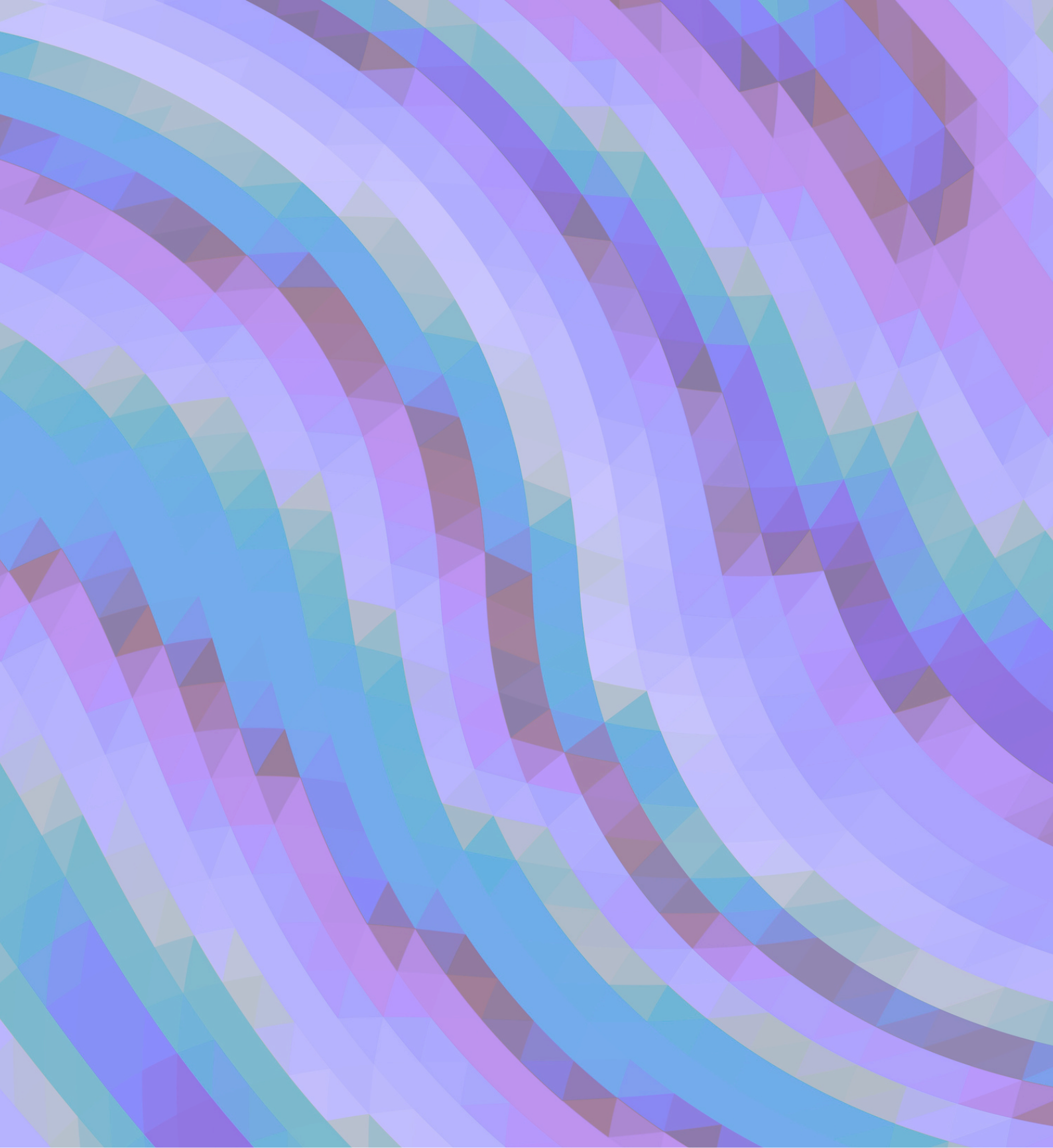
РЕКОМЕНДАЦИИ

КОМПАНИЯМ, ВЛАДЕЮЩИМ СОЦИАЛЬНЫМИ СЕТЯМИ:

- Обеспечить строгое соответствие **модерации и курирования контента** международно-правовым нормам в сфере защиты прав человека и не препятствовать реализации права защищать права человека;
- Воздерживаться от создания механизмов на основе **ИИ** и их использования для модерации и удаления контента без независимого фактчекинга и верификации людьми;
- Повысить **прозрачность** решений по модерации контента и процесса обжалования;
- Усовершенствовать инструменты для обеспечения приватности и безопасности, включая более надежное шифрование, возможность направления анонимных жалоб и защиту от доксинга и слежки;
- Провести оценку **эффективности механизмов для сообщений** об онлайн-харассменте и выделить ресурсы для создания доверительных, доступных механизмов **с участием живых людей** для оказания помощи нуждающимся в ней правозащитникам с привлечением сотрудников, обученных соответствующим языкам, культурным особенностям и политическим контекстам;
- Сократить сроки ожидания ответа, особенно **для доверенных партнеров и в экстренных случаях**;
- Выделить ресурсы на создание инструментов и протоколов модерации **с учетом гендерного фактора** и включить отдельные разделы о гендерном харассменте и насилии в правила сообщества;
- Провести независимые **проверки алгоритмов**, чтобы прекратить распространение дезинформации и контента с языком ненависти, а также взять на себя обязательства по разработке инклюзивных алгоритмов, которые не маргинализируют правозащитный контент;
- Отказываться от **выполнения требований государств**, противоречащих международно-правовым нормам в сфере защиты прав человека;
- Публично фиксировать и **публиковать требования государств об удалении** контента и ответы на них платформ;
- Отменить недавние решения об **отказе от политик в области языка ненависти, дезинформации и приватности**, а также вернуть сотрудников в программы по обеспечению доверия и безопасности и по правам человека;
- Создавать и расширять **каналы для коммуникаций с правозащитниками и гражданским обществом**, особенно в регионах, где для этого не существует структурных механизмов, а также обеспечить их участие в разработке эффективных с их точки зрения политик и систем поддержки.

ГОСУДАРСТВАМ:

- Обеспечивать соблюдение компаниями, владеющими социальными сетями, принципа свободы выражения мнений и **поддержку ими стандартов по защите прав человека**;
- Разрабатывать **регуляторные решения**, которые могли бы решать проблему доминирования основных платформ;
- Признать понятие цифрового насилия как одного из видов широкого круга действий, направленных против правозащитников;
- Расследовать случаи онлайн-насилия в отношении правозащитников и привлекать виновных к ответственности;
- Разработать и реализовывать четкие **законодательные и программные принципы**, предусматривающие ответственность платформ социальных сетей за создание условий для онлайн-харассмента, особенно когда такие действия угрожают безопасности правозащитников;
- Воздерживаться от принуждения социальных сетей к осуществлению цензуры или слежки за пользователями.



HRC-SR-DEFENDERS@UN.ORG
WWW.SRDEFENDERS.ORG



СПЕЦИАЛЬНЫЙ
ДОКЛАДЧИК
ООН ПО
ПОЛОЖЕНИЮ
ПРАВООЩИТНИКОВ

С СОЛИДАРНОСТЬЮ И НАДЕЖДОЙ