

**Совет по правам человека****Тридцать восьмая сессия**

18 июня – 6 июля 2018 года

Пункт 3 повестки дня

**Поощрение и защита всех прав человека,
гражданских, политических, экономических,
социальных и культурных прав,
включая право на развитие****Доклад Специального докладчика по вопросу
о поощрении и защите права на свободу мнений
и их свободное выражение****Записка секретариата**

Секретариат имеет честь препроводить Совету по правам человека доклад Специального докладчика по вопросу о поощрении и защите права на свободу мнений и их свободное выражения Дэвида Кэя, подготовленный в соответствии с резолюцией Совета 34/18. В своем докладе Специальный докладчик рассматривает вопросы регулирования пользовательского интернет-контента. Он рекомендует государствам обеспечивать формирование благоприятных условий для свободы выражения мнений в Интернете, а компаниям рекомендует применять правозащитные стандарты на всех этапах своих операций. Право прав человека предоставляет компаниям средства для формулирования своей позиции с соблюдением демократических норм и воспрепятствования авторитарным требованиям. Компании и государства должны как минимум стремиться к выходу на принципиально более высокий уровень прозрачности, начиная уже с этапа разработки правил и вплоть до их применения, в интересах обеспечения автономии пользователей в условиях все более широкого осуществления отдельными лицами своих прав в Интернете.



**Доклад Специального докладчика по вопросу
о поощрении и защите права на свободу мнений
и их свободное выражение**

Содержание

	<i>Стр.</i>
I. Введение	3
II. Правовая база.....	4
A. Обязательства государств	4
B. Обязанности компаний.....	5
III. Основные обеспокоенности, связанные с регулированием контента	7
A. Государственное регулирование	7
B. Модерация контента компаниями	10
IV. Правозащитные принципы в контексте модерации контента силами компаний	17
A. Основные стандарты модерации контента	18
B. Процессы модерации в компаниях и связанная с ними деятельность	20
V. Рекомендации	24

I. Введение

1. На заре цифровой эры Джон Перри Барлоу заявил, что Интернет создает «мир, где любой и везде может выражать его или ее убеждения независимо от того, насколько они необычны, без страха быть принужденным к молчанию или к конформизму»¹. И хотя Интернет остается величайшим в истории инструментом всеобщего доступа к информации, сегодня уже трудно встретить подобные проповеди в виртуальном пространстве. Общество находит в пользовательском контенте проявления ненависти, оскорбления и дезинформацию. Правительства же усматривают вербовку террористов или вызывающие у них обеспокоенность элементы инакомыслия или оппозиции. Организации гражданского общества отмечают, что неподотчетным частным субъектам передаются функции государства, в частности по защите свободы выражения мнений. Предпринимая определенные шаги по разъяснению своих правил и механизмов взаимодействия с правительствами, компании, тем не менее, остаются загадочными регуляторами, формирующими некое «платформенное право» с расплывчатыми критериями прозрачности, согласованности, подотчетности и правовой защиты. Организация Объединенных Наций, региональные организации и договорные органы подтверждают, что осуществляемые в реальной жизни права равным образом применимы и в виртуальном пространстве, однако не всегда понятно, действительно ли компании сами защищают права своих пользователей или же государства побуждают их к этому методами правового стимулирования.

2. В настоящем докладе Специальный докладчик предлагает рамочную схему модерации пользовательского онлайн-контента, основанную на соблюдении прав человека². Он пытается дать ответы на коренные вопросы: каковы обязанности компаний по недопущению ущемления их платформами прав, гарантируемых по международному праву? Какие стандарты они должны применять по отношению к модерации контента? Должны ли государства регулировать модерацию коммерческого контента и, если да, то каким образом? Право предполагает прозрачность и подотчетность государств в их усилиях по ослаблению угроз для свободы выражения мнений. Должны ли мы этого ожидать и от частных субъектов? Что представляют собой процессы защиты и правовой помощи в цифровую эпоху?

3. Некоторые из этих вопросов были рассмотрены в предыдущих докладах³. В настоящем докладе основное внимание уделяется регулированию пользовательского контента в первую очередь государствами и операторами социальных сетей, причем такими методами, которые применимы ко всем соответствующим субъектам в секторе информационно-коммуникационных технологий (ИКТ). Специальный докладчик излагает основные принципы применимой правозащитной нормативной базы и подходы компаний и государств к регулированию контента. Он предлагает стандарты и процессы, которые надлежит применять компаниям для регулирования контента в соответствии с правом прав человека.

4. Первоначальной основой для подготовки настоящего доклада стали результаты изучения условий предоставления услуг компаниями, отчетности об обеспечении прозрачности и вспомогательных источников информации. На предложение сообщить замечания от государств поступил 21 ответ и 29 ответов прислали негосударственные субъекты (в том числе 1 ответ от компаний). В стремлении понять их подходы к модерации контента Специальный докладчик посетил несколько компаний в Силиконовой долине и провел беседы с другими субъектами⁴. В 2017 и

¹ John Perry Barlow, A Declaration of the Independence of Cyberspace, 8 February 1996.

² Под термином «модерация» понимается процесс, с помощью которого интернет-компания определяют, соответствует ли пользовательский контент стандартам, закрепленным в их условиях предоставления услуг и в других правилах.

³ A/HRC/35/22 и A/HRC/32/38.

⁴ Специальный докладчик посетил штаб-квартиры «Фейсбука», «ГитХаба», «Гугла», «Реддита» и «Твиттера» и провел беседы с представителями «Яху/Оат», «Лайн» и «Майкрософт». Он

2018 годах в Бангкоке и Женеве он провел полезные консультации с представителями гражданского общества, а в 2018 году были организованы обсуждения в онлайн-режиме с экспертами из Латинской Америки, Ближнего Востока и Северной Африки, а также из стран Африки, расположенных к югу от Сахары⁵.

II. Правовая база

5. Деятельность компаний в секторе ИКТ среди прочего подразумевает права на неприкосновенность частной жизни, свободу религии и убеждений, мнений и их выражения, собраний и ассоциации, а также на право на участие в общественной жизни. Настоящий доклад посвящен свободе выражения мнений, и в нем признается взаимозависимость прав, в частности важное значение неприкосновенности частной жизни как залога свободы выражения мнений⁶. В статье 19 Международного пакта о гражданских и политических правах зафиксированы ратифицированные 170 государствами и отражающие Всеобщую декларацию прав человека универсально принятые нормы, которые гарантируют «право беспрепятственно придерживаться своих мнений» и «право искать, получать и распространять всякого рода информацию и идеи, независимо от государственных границ» и любыми способами⁷.

A. Обязательства государств

6. Право прав человека налагает на государства обязанности по обеспечению благоприятных условий для свободы выражения мнений и по защите ее осуществления. Обязанность обеспечивать свободу выражения мнений среди прочего требует от государства поощрять разнообразие и независимость средств массовой информации и доступ к информации⁸. Кроме того, международные и региональные органы настоятельно призывают государства содействовать универсальному доступу к Интернету⁹. На государства также возложена обязанность обеспечивать невмешательство частных субъектов в осуществление свободы мнений и свободы их выражения¹⁰. В Руководящих принципах по вопросам предпринимательской деятельности в аспекте прав человека, принятых Советом по правам человека в 2011 году, особое внимание уделяется принципу 3, касающемуся обязанностей государств обеспечивать создание условий, способствующих соблюдению предприятиями прав человека¹¹.

7. Государства не могут ограничивать право беспрепятственно придерживаться мнений. Согласно пункту 3 статьи 19 Пакта вводимые государствами ограничения свободы выражения мнений должны соответствовать следующим твердо установленным критериям:

также посетил некоммерческий фонд «Викимедия». В рамках связанной с настоящим докладом работы он предполагает посетить компании в Москве, Пекине, Сеуле и Токио.

⁵ Специальный докладчик хотел бы выразить признательность своему юрисконсульту Амосу Тохе и студентам Лаборатории международного правосудия юридического факультета Калифорнийского университета в Ирвайне.

⁶ См. A/HRC/29/32, пункты 16–18.

⁷ См. также African Charter on Human and Peoples' Rights, art. 9; American Convention on Human Rights, art. 13; Convention for the Protection of Human Rights and Fundamental Freedoms, art. 10. См. также Centro de Estudios en Libertad de Expresión y Acceso a la Información submission.

⁸ Совместное заявление о свободе выражения мнений и «фейковых новостях», дезинформации и пропаганде, 3 марта 2017 года, раздел 3. См. также Комитет по правам человека, замечание общего порядка № 34 (2011) о свободе мнений и их выражения, пункты 18 и 40; A/HRC/29/32, пункт 61 и A/HRC/32/38, пункт 86.

⁹ См. Совет по правам человека, резолюция 32/13, пункт 12; Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights, *Standards for a Free, Open and Inclusive Internet* (2016), para. 18.

¹⁰ См. замечание общего порядка № 34, пункт 7.

¹¹ A/HRC/17/31.

- *Законность*. Ограничения должны быть «установлены законом». В частности, они должны быть приняты в рамках стандартных юридических процедур и ограничивать дискреционные полномочия правительства таким образом, который позволяет с «достаточной точностью» установить различие между законным и незаконным выражением мнений. Этому основополагающему требованию не удовлетворяют принимаемые в тайне ограничения¹². Гарантия законности в целом должна предусматривать контроль со стороны независимых судебных органов¹³.
- *Необходимость и соразмерность*. Государства должны демонстрировать, что вводимое ограничение налагает минимальное бремя на осуществление того или иного права и в данном случае фактически защищает или может защищать законные интересы государства. Государства не могут ограничиваться заявлениями о наличии необходимости и должны подтверждать ее при принятии ограничительного законодательства и ограничении свободного выражения конкретного мнения¹⁴.
- *Легитимность*. Для того чтобы любое ограничение было законным, оно должно защищать только те интересы, которые перечислены в пункте 3 статьи 19: права и репутацию других лиц, государственную безопасность или общественный порядок, либо здоровье или нравственность населения. Ограничения, предназначенные для защиты прав других лиц, включают, например, «права человека, признанные в Пакте и, более широко, в международном законодательстве в области прав человека»¹⁵. Ограничения в целях защиты, в частности, прав на неприкосновенность личной жизни, на жизнь, на процессуальные гарантии, ассоциацию или участие в общественной жизни будут легитимными в случае подтверждения того, что они удовлетворяют критериям легитимности и необходимости. В стремлении обеспечить, чтобы вводимые ограничения отражали принципы недискриминации и универсальности прав, Комитет по правам человека предостерегает, что ограничения с целью защиты «нравственности населения» не должны основываться «на принципах, вытекающих исключительно из одной единственной традиции»¹⁶.

8. Предусмотренные пунктом 2 статьи 20 Пакта ограничения, требующие от государств запрета «всякого выступления в пользу национальной, расовой или религиозной ненависти, представляющего собой подстрекательство к дискриминации, вражде или насилию», также должны удовлетворять совокупным критериям законности, необходимости и легитимности¹⁷.

В. Обязанности компаний

9. Интернет-компании стали центральными площадками для обсуждений и дискуссий, доступа к информации, торговли и развития человеческого потенциала¹⁸. Они осуществляют сбор и хранение личных данных о миллиардах отдельных лиц, в том числе сведения об их привычках, местонахождении и деятельности, и часто претендуют на выполнение какой-либо гражданской роли. В 2004 году «Гугл» заявил о своем благородном стремлении делать «для мира добрые дела, даже если нам придется отказаться от некоторых краткосрочных выгод»¹⁹. Основатель «Фейсбука»

¹² Там же, пункт 25; A/HRC/29/32.

¹³ Там же.

¹⁴ См. замечание общего порядка № 34, пункт 27.

¹⁵ Там же, пункт 28.

¹⁶ Там же, пункт 32.

¹⁷ Там же, пункт 50. См. также A/67/357.

¹⁸ См., например, Supreme Court of the United States, *Packingham v. North Carolina*, opinion of 19 June 2017; European Court of Human Rights, *Times Newspapers Ltd. (Nos. 1 and 2) v. The United Kingdom* (application Nos. 3002/03 and 23676/03), judgment of 10 March 2009, para. 27.

¹⁹ Securities Registration Statement (S-1) under the Securities Act of 1933, 18 August 2004.

провозгласил желание «развивать социальную инфраструктуру, чтобы дать людям силы для построения глобального сообщества, работающего в наших общих интересах»²⁰. «Твиттер» обещает проводить политику, которая направлена на «совершенствование свободного и глобального диалога, и не отвлекает нас от него»²¹. Российская социальная сеть «В контакте» «объединяет людей всего мира», а «Тенсент» придерживается стилистики правительства Китая, заявляя, что ставит перед собой цель «помочь в построении гармоничного общества и стать добропорядочным корпоративным гражданином»²².

10. Принципы защиты прав человека в своей деятельности применяют лишь немногие компании, причем, по мнению большинства из них, они сводятся к методам реагирования на угрозы и требования со стороны правительств²³. Однако в Руководящих принципах по вопросам предпринимательской деятельности в аспекте прав человека устанавливается «всеобщая норма[ы] ожидаемого поведения», которая должна применяться в рамках всех операций компаний, где бы они не осуществляли свою деятельность²⁴. Руководящие принципы не являются обязывающими, однако огромная роль компаний в общественной жизни в целом настоятельно требует их принятия и соблюдения.

11. Руководящие принципы закладывают рамочную основу, согласно которой компании как минимум должны:

а) в своей деятельности избегать оказания или содействия оказанию неблагоприятного воздействия на права человека и стремиться предотвращать или смягчать неблагоприятное воздействие, которое непосредственно связано с их деятельностью, продукцией или услугами вследствие их деловых отношений, даже если они непосредственно не способствовали оказанию такого воздействия (принцип 13);

б) взять на себя обязательства соблюдать права человека своих пользователей в заявлении программного характера (принцип 16);

в) осуществлять должную заботу о правах человека с целью выявления, предотвращения и представления отчетности о фактическом и потенциальном воздействии на права человека их деятельности на основе регулярной оценки рисков и воздействия, конструктивных консультаций с потенциально затрагиваемыми группами и другими заинтересованными сторонами, а также надлежащих последующих действий по смягчению или предотвращению такого воздействия (принципы 17–19);

г) проводить стратегии по предупреждению и смягчению воздействия, которые в максимально возможной степени соблюдают принципы международно признанных прав человека, когда они сталкиваются с противоречивыми требованиями местных законов (принцип 23);

д) постоянно отслеживать свои усилия по соблюдению прав, в том числе посредством регулярных консультаций с заинтересованными сторонами и поддержания разветвленных, доступных и эффективных связей с затрагиваемыми группами и общественностью (принципы 20–21);

е) обеспечивать надлежащее устранение нарушений, в том числе через оперативные механизмы рассмотрения жалоб, к которым пользователи могут получить доступ без обострения у них «чувства собственного бессилия» (принципы 22, 29 и 31).

²⁰ Mark Zuckerberg, “Building global community”, Facebook, 16 February 2017.

²¹ Twitter, S-1 Registration Statement, 13 October 2013, pp. 91–92.

²² «В контакте», информация о компании; Tencent, “About Tencent”.

²³ Материал представлен Датским институтом по правам человека. Материал представлен Cf. Yahoo/Oath, 2016 год.

²⁴ Руководящие принципы, принцип 11.

III. Основные обеспокоенности, связанные с регулированием контента

12. Правительства стремятся устанавливать условия модерирования контента компаниями, а компании обуславливают индивидуальный доступ к своим платформам согласием пользователей с условиями предоставления услуг, в которых определяется, какие мнения и какими способами могут выражать отдельные лица.

A. Государственное регулирование

13. Государства регулярно требуют от компаний четко ограничивать такой незаконный контент, как демонстрация сексуальных надругательств над детьми, прямые и достоверные угрозы причинения вреда либо подстрекательство к насилию, при том условии, что они также отвечают критериям законности и необходимости²⁵. Некоторые государства идут еще дальше и в интересах формирования онлайн-регулятивной среды вводят цензуру и уголовную ответственность²⁶. Широко сформулированные ограничительные законы об «экстремизме», богохульстве, клевете, «оскорбительных» высказываниях, «фейковых новостях» и «пропаганде» часто используются как предлог для предъявления компаниям требований о блокировке легитимного дискурса²⁷. Государства все чаще принимают целенаправленные меры именно в отношении контента интернет-платформ²⁸. Прочие законы могут оказывать на конфиденциальность в сети такое воздействие, которое препятствует осуществлению свободы мнений и их выражения²⁹. Помимо этого, в целях ограничения доступности независимых сетей и ослабления доверия к ним многие государства применяют методы дезинформации и пропаганды³⁰.

14. *Механизмы защиты от ответственности.* Уже на заре цифровой эры многие государства вводили правила защиты посредников от ответственности за размещаемый третьими сторонами на их платформах контент. Так, например, в соответствии с директивой Европейского союза об интернет-торговле устанавливается правовой режим защиты посредников от ответственности за контент за исключением тех случаев, когда они выходят за рамки своей роли «простых проводников», «хранилищ» или «площадок для размещения» пользовательской информации³¹. Статья 230 Закона Соединенных Штатов Америки о благопристойности средств связи в целом предусматривает иммунитет провайдеров «интерактивных компьютерных услуг», которые размещают или публикуют информацию о других сторонах, однако сфера ее действия применения уже ограничена³². Действующий в Бразилии режим

²⁵ В Ирландии создаются совместные с компаниями механизмы регулирования контента в целях ограничения незаконных материалов о сексуальных надругательствах над детьми: материал представлен Ирландией. Многие компании используют алгоритм распознавания изображений для выявления и удаления детской порнографии: материалы представлены Открытым технологическим институтом, стр. 2 и организацией «Статья 19», стр. 8.

²⁶ См. A/HRC/32/38, пункты 46–47. Информацию по вопросам отключения Интернета см. A/HRC/35/22, пункты 8–16 и примеры сообщений Специального докладчика: Nos.UA TGO 1/2017, UA IND 7/2017 и AL GMB 1/2017.

²⁷ Сообщения № OL MYS 1/2018; UA RUS 7/2017; UA ARE 7/2017, AL BHR 8/2016, AL SGP 5/2016 и OL RUS 7/2016. В Азербайджане запрещена пропаганда терроризма, религиозного экстремизма и суицидов: материал представлен Азербайджаном.

²⁸ См. сообщения № OL PAK 8/2016 и OL LAO 1/2014; Association for Progressive Communications *Unshackling Expression: A Study on Laws Criminalising Expression Online in Asia*, GISWatch 2017 Special Edition.

²⁹ A/HRC/29/32.

³⁰ См., например, Gary King, Jennifer Pan and Margaret E. Roberts, "How the Chinese Government fabricates social media posts for strategic distraction, not engaged argument", *American Political Science Review*, vol. 111, No. 3 (2017), pp. 484–501.

³¹ Directive No. 2000/31/EC of the European Parliament and of the Council of 8 June 2000.

³² 47 United States Code § 230. See also the Allow States and Victims to Fight Online Sex Trafficking Act (H.R. 1865).

ответственности посредников требует получения судебного постановления об ограничении того или иного конкретного контента³³, а аналогичный режим в Индии предусматривает процесс «уведомления и закрытия интернет-ресурса», запускаемый по решению суда или аналогичного органа юстиции³⁴. В разработанных в 2014 году коалицией экспертов гражданского общества Манильских принципах ответственности посредников определяются основные принципы, на которые должны опираться любые нормативные положения об ответственности посредников.

15. *Возложение обязательств на компании.* Некоторые государства возлагают на компании обязательства по ограничению контента на основании размытых или сложных юридических критериев без предварительного судебного разбирательства и под угрозой строгих мер наказания. Так, например, Закон Китая о кибербезопасности 2016 года усиливает нечетко сформулированные запреты на распространение «ложной» информации, способной подорвать «общественный или экономический уклад», национальное единство или национальную безопасность; он также требует от компаний отслеживать деятельность своих сетей и сообщать о нарушениях властям³⁵. Сообщалось, что за невыполнение этих требований крупнейшие социальные сети страны были наказаны крупными штрафами³⁶.

16. Обязательства отслеживать и быстро удалять пользовательский контент приобретают также глобальные масштабы, способствуя формированию режима санкций, способного подорвать свободу выражения мнений даже в демократических обществах. В Германии Закон о контроле над социальными сетями («NetzDG») требует от крупных операторов социальных сетей удалять контент, который не соответствует оговоренным местным законам, и предусматривает за его несоблюдение применение в предельно короткие сроки серьезных мер наказания³⁷. Европейская комиссия даже рекомендовала государствам-членам ввести юридические обязательства по активному мониторингу и фильтрации незаконного контента³⁸. В соответствии с принятыми в Кении в 2017 году руководящими принципами распространения контента в социальных сетях во время выборов интернет-платформы обязаны в течение 24 часов «закрывать аккаунты, используемые для распространения на их платформах нежелательного политического контента»³⁹.

17. В свете законных обеспокоенностей государств, в частности вопросами неприкосновенности частной жизни и национальной безопасности, становится вполне понятна привлекательность регулирования контента. Однако правила связаны с рисками для свободы выражения мнений и оказывают столь значительное давление на компании, что они могут удалять законный контент в общем стремлении избежать ответственности. Они также предусматривают делегирование регулятивных функций частным субъектам, которые не располагают основными инструментами обеспечения подотчетности. Требования быстрого и автоматического удаления контента могут стать причиной появления новых форм предварительного ограничения, которые уже и так угрожают творческой работе в контексте авторского права⁴⁰. Сложные вопросы реально существующего положения дел и законодательства, как правило, должны

³³ *Marco Civil da Internet*, federal law 12.965, arts. 18–19.

³⁴ Supreme Court of India, *Shreya Singhal v. Union of India*, decision of 24 March 2015.

³⁵ Articles 12 and 47; Human Rights in China submission, 2016, p. 12. For comments on an earlier draft of the Cybersecurity Law, see communication No. OL CHN 7/2015. См. также Global Voices, “Netizen Report: Internet censorship bill looms large over Egypt”, 16 March 2018; Republic of South Africa, Films and Publications Amendment Bill (B 61—2003).

³⁶ PEN America, *Forbidden Feeds: Government Controls on Social Media in China* (2018), p. 21.

³⁷ Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), July 2017. См. сообщение № OL DEU 1/2017.

³⁸ European Commission, recommendation on measures to effectively tackle illegal content online (last updated: 5 March 2018).

³⁹ См. сообщение № OL KEN 10/2017; Javier Pallero, “Honduras: new bill threatens to curb online speech”, Access Now, 12 February 2018.

⁴⁰ См. European Commission, Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, COM (2016) 593 final, art. 13; Daphne Keller, “Problems with filters in the European Commission’s platforms proposal”, Stanford Law School Center for Internet and Society, 5 October 2017; Fundación Karisma submission, 2016, pp. 4–6.

рассматриваться государственными структурами, а не частными субъектами, которые могут применять несовместимые с надлежащими процессуальными нормами процедуры и руководствоваться главным образом экономическими соображениями⁴¹.

18. *Случаи удаления контента в глобальном масштабе.* Некоторые государства требуют экстерриториального удаления ссылок, веб-сайтов и другого контента, который предположительно нарушает местные законы⁴². Такие требования вызывают серьезную обеспокоенность в связи с возможностью ущемления государствами права на свободу выражения мнений «без учета национальных границ». Сама логика таких требований допускает трансграничную цензуру в угоду наиболее консервативным цензорам. Стороны, требующие удаления контента, должны подавать надлежащие заявки в каждой соответствующей юрисдикции в рамках стандартных юридических и судебных процедур.

19. *Требования правительств, не основанные на национальном законодательстве.* Компании проводят различие между требованиями об удалении предположительно незаконного контента, поступающими по обычным предусмотренным законодательством каналам, и требованиями удаления контента на основании условий предоставления услуг, которые устанавливают сами компании⁴³. (Законные требования об удалении контента обычно удовлетворяются только в запрашивающей юрисдикции; удаление контента на основании условий предоставления услуг, как правило, осуществляется в глобальном масштабе.) Государственные органы все чаще стараются добиться удаления контента вне рамок правового процесса или даже путем предъявления требований на основании условий предоставления услуг⁴⁴. В некоторых странах уже созданы специальные правительственные структуры, направляющие в компании требования об удалении контента. Так, например, Группа по контролю Интернета Европейского союза «выявляет в Интернете террористический и насильственный экстремистский контент и сотрудничает с провайдерами интернет-услуг с целью его удаления»⁴⁵. Аналогичный механизм существует и в Австралии⁴⁶. Сообщается, что в Юго-Восточной Азии связанные с правительствами стороны якобы пытаются использовать условия предоставления услуг для ограничения политической критики⁴⁷.

20. Государство также оказывает на компании давление в целях более оперативного удаления ими контента посредством необязующих механизмов, большинству из которых свойственна ограниченная транспарентность. В результате введения в Пакистане трехлетнего запрета на «Ютьюб» «Гугл» был вынужден создать местную версию в ответ на требования правительства об удалении «оскорбительного» контента⁴⁸. «Фейсбук» и Израиль предположительно договорились координировать

⁴¹ В соответствии с законодательством Европейского союза поисковые системы обязаны проверять обоснованность жалоб, подаваемых в порядке реализации «права быть забытым». European Court of Justice, *Google Spain v. Agencia Española de Protección de Datos and Mario Costeja González* (case C-131/12), judgment (Grand Chamber) of 13 May 2014; submissions by ARTICLE 19, pp. 2–3 and Access Now, pp. 6–7; Google, “Updating our ‘right to be forgotten’ Transparency Report”; Theo. Bertram and others, *Three Years of the Right to be Forgotten* (Google, 2018).

⁴² См., например, PEN America, *Forbidden Feeds*, pp. 36–37; Supreme Court of Canada, *Google Inc. v. Equustek Solutions Inc.*, judgment of 28 June 2017; European Court of Justice, *Google Inc. v. Commission nationale de l’informatique et des libertés (CNIL)* (case C-507/17); первоначальный материал представлен организацией «Глобальная сетевая инициатива», стр. 6.

⁴³ Сопоставьте Twitter Transparency Report: Removal Requests (January–June 2017) с Twitter Transparency Report: Government Terms of Service Reports (January–June 2017). См. также Facebook, Government requests: Frequently Asked Questions (FAQs).

⁴⁴ Материалы представлены организациями «Статья 19», стр. 2 и «Глобальная сетевая инициатива», стр. 5.

⁴⁵ European Union, Internet Referral Unit, Year One Report, sect. 4.11; материалы представлены Европейской организацией за цифровые права (EDRi), стр. 1, и организацией «Экссес нау», стр. 2–3.

⁴⁶ Материал представлен Австралией.

⁴⁷ Альянс печатных СМИ Юго-Восточной Азии, стр. 1.

⁴⁸ Материал представлен Фондом цифровых прав.

усилия и согласовывать персонал по вопросам мониторинга и удаления «подстрекательств» в онлайн-режиме. Детали этой договоренности не разглашаются, однако Министр юстиции Израиля утверждал, что в период с июня по сентябрь 2016 года «Фейсбук» выполнил практически все просьбы правительства об удалении «подстрекательств»⁴⁹. Существование договоренностей о координации действий по удалению контента с участием государства усиливает обеспокоенность по поводу выполнения компаниями общественных функций без надзора со стороны судов или других механизмов обеспечения подотчетности⁵⁰.

21. Кодекс поведения Европейского союза 2016 года по противодействию незаконным ненавистническим высказываниям в Интернете предусматривает заключение соглашения об удалении контента между Европейским союзом и четырьмя основными компаниями, обязующего их сотрудничать с «доверенными модераторами контента» и поощрять распространение независимыми пользователями контрпропагандистских материалов⁵¹. И если распространение контрпропагандистских аргументов может быть целесообразным перед лицом «экстремистского» или «террористического» контента, то давление с целью принуждения к применению таких подходов сопряжено с риском превращения интернет-платформ в носителей пропаганды за рамками определенных областей законной обеспокоенности⁵².

В. Модерация контента компаниями

Соблюдение компаниями национального законодательства

22. Каждая компания придерживается принципа соблюдения законов того места, в котором она осуществляет свои операции. Как указывалось в «Фейсбуке»: «Если после тщательной юридической проверки мы приходим к выводу, что по местному законодательству контент является незаконным, мы удаляем его в соответствующей стране или на соответствующей территории»⁵³. «Тенсент», владелец мобильного чата и используемого в социальных сетях приложения «ВиЧет», идет значительно дальше и требует от любого пользователя платформы в Китае и от китайских граждан, использующих эту платформу «в любом районе мира», соблюдения ограничений по контенту, отражающих политику или законодательство Китая⁵⁴. Несколько компаний также сотрудничают друг с другом и с надзорными органами в интересах удаления изображений сексуальных надругательств над детьми⁵⁵.

23. Выполнение обязательства по соблюдению законов может быть затруднено в тех случаях, когда соответствующее государственное законодательство сформулировано нечетко, допускает расхождения в толковании или не соответствует праву прав человека. Так, например, законы о борьбе с «экстремизмом», содержащие нечеткое определение ключевого термина, предоставляют государственным органам свободу действий для оказания на компании давления в интересах удаления контента

⁴⁹ «7amleh» – материал, представленный Арабским центром развития социальных сетей.

⁵⁰ Ассоциация поддержки прогрессивных коммуникаций, стр. 14 и «7amleh».

⁵¹ «Выражение "доверенные модераторы контента" касается статуса, который предоставляется определенным организациям и позволяет им сообщать о незаконном контенте через посредство специальной системы или канала уведомления, которые недоступны для обычных пользователей». European Commission, Code of Conduct on countering illegal hate speech online: First results on implementation (December 2016).

⁵² В стремлении к выработке общепромышленных технологических инструментов удаления террористического контента со своих сайтов эти же компании организовали Глобальный интернет-форум по противодействию терроризму. Google, "Update on the Global Internet Forum to Counter Terrorism", 4 December 2017.

⁵³ Facebook, Government requests: FAQs. См. также Google legal removal requests; Twitter rules and policies; Reddit content policy.

⁵⁴ Tencent, Terms of Service: Introduction; Tencent, Agreement on Software License and Service of Tencent Wenxin.

⁵⁵ Организация Объединенных Наций по вопросам образования, науки и культуры, *Fostering Freedom Online: The Role of Internet Intermediaries* (Paris, 2014), pp. 56–57.

по сомнительным причинам⁵⁶. Аналогичным образом на компании нередко оказывают давление с целью принуждения их к соблюдению государственных законов, предусматривающих уголовную ответственность за контент, который якобы является, например, богохульственным, содержит критику государства, клевету на госчиновников или носит недостоверный характер. Как поясняется ниже, Руководящие принципы предусматривают инструменты минимизации воздействия таких законов на отдельных пользователей. Глобальная сетевая инициатива, в которой участвуют многие заинтересованные стороны и которая помогает ИКТ-компаниям решать проблемы прав человека, разработала дополнительные инструктивные материалы по методике использования таких инструментов⁵⁷. Одним из таких инструментов является транспарентность; многие компании представляют ежегодные отчеты о количестве полученных от правительств и выполненных ими требований в расчете на одно государство⁵⁸. Однако компании не всегда раскрывают достаточно сведений о принятых по требованиям правительства мерах и регулярно не отчитываются о предъявляемых правительствами требованиях в соответствии с условиями предоставления услуг⁵⁹.

Стандарты модерации, принятые компаниями

24. Интернет-компании требуют от своих пользователей соблюдать условия предоставления услуг и «стандарты сообщества», которые регулируют выражение мнений на их платформах⁶⁰. Вводимые компаниями условия предоставления услуг, которые пользователи обязаны принимать в обмен на использование их платформ, содержат указания о юрисдикциях для разрешения споров и сохраняют за компаниями право на принятие по своему усмотрению любых мер в отношении контента и аккаунтов⁶¹. Политика платформ в отношении контента входит в число таких условий и предусматривает ограничения мнений, которые могут выражать пользователи, и способов их выражения. Большинство компаний не разрабатывают стандарты контента непосредственно на базе какого-либо конкретного свода законов, которые могут регулировать выражение мнений, в частности на национальном праве или международном праве прав человека. Вместе с тем гигантский китайский поисковик «Байду» запрещает контент, «противоречащий основным принципам Конституции» Китайской Народной Республики⁶².

25. К разработке политики по вопросам модерации контента обычно привлекаются юристы, менеджеры по вопросам публичной политики и программным продуктам, а также старшие должностные лица. В целях борьбы со спамом, мошенничеством и злоупотреблениями компании могут создавать «группы по вопросам доверия и безопасности», а группы по противодействию терроризму могут регулировать

⁵⁶ См. Maria Kravchenko, “Inappropriate enforcement of anti-extremist legislation in Russia in 2016”, SOVA Center for Information and Analysis, 21 April 2017; Danielle Citron, “Extremist speech, compelled conformity, and censorship creep”, *Notre Dame Law Review*, vol. 93, No. 3 (2018), pp. 1035–1071.

⁵⁷ Global Network Initiative, Principles on Freedom of Expression and Privacy, sect. 2. Social media companies participating in the Initiative include Facebook, Google, Microsoft/LinkedIn and Yahoo/Oath.

⁵⁸ См. пункт 39 ниже. Кроме того, «Аутоматтик», «Гугл», «Майкрософт»/«Бинг» и «Твиттер» входят в число компаний, которые регулярно, хотя необязательно в полном объеме, размещают в базе данных «Люмен» (Lumen database) информацию о требованиях правительств по вопросам закрытия аккаунтов и интеллектуальной собственности.

⁵⁹ Ranking Digital Rights, 2017 Corporate Accountability Index, p. 28.

⁶⁰ Jamila Venturini and others, *Terms of Service and Human Rights: An Analysis of Online Platform Contracts* (Rio de Janeiro, Revan, 2016).

⁶¹ Baidu user agreement (“[We] remove and delete any content in this service based on Baidu’s own discretion for any reason.”); Tencent terms of service (“We reserve the right to block or remove Your Content for any reason, including as is in our opinion appropriate or as required by applicable laws and regulations.”); Twitter terms of service (“We may suspend or terminate your account or cease providing you with all or part of the Services at any time for any or no reason.”).

⁶² Baidu terms of service, sect. 3.1.

террористический контент⁶³. Некоторые компании создают механизмы для стимулирования внешних групп к представлению замечаний по узкоспециализированным аспектам политики в отношении контента⁶⁴. Многократное увеличение пользовательского контента стимулировало разработку детальных и постоянно корректируемых правил. Такие правила меняются в зависимости от целого ряда факторов, от размера, доходов и бизнес-модели компании до «бренда и репутации платформы, ее устойчивости к рискам и типа взаимодействия с пользователями, которых она стремится привлечь»⁶⁵.

Вопросы, вызывающие обеспокоенность в связи со стандартами контента

26. *Нечеткие правила.* Вводимые компаниями запреты на угрозы или пропаганду терроризма⁶⁶, поддержку или восхваление руководителей опасных организаций⁶⁷, а также на контент, способствующий совершению террористических актов и разжиганию насилия⁶⁸, подобно контртеррористическому законодательству сформулированы слишком нечетко⁶⁹. В политике компаний в отношении ненависти, притеснений и злоупотреблений также не содержится четкой квалификации преступления. Установленный «Твиттером» запрет на «поведение, вызывающее страх перед какой-либо защищенной группой», равно как и проводимое «Фейсбуком» различие между «прямыми нападениями» на защищаемые особенности или попросту «обидным или оскорбительным контентом», являются субъективной и непрочной основой для модерации контента⁷⁰.

27. *Ненависть, притеснения, злоупотребления.* Нечеткость политики в отношении ненавистнических высказываний и притеснений стала причиной жалоб на ее непоследовательное проведение в ущерб меньшинствам при одновременном укреплении статуса доминантных или влиятельных групп. Пользователи и представители гражданского общества сообщают о случаях насилия и надругательств над женщинами, включая угрозы физического насилия, женоненавистнические высказывания, размещение в Интернете без получения на это согласия интимных или фейковых изображений, а также доксинг⁷¹; угрозы причинения вреда в адрес лишенных политических прав групп⁷², расовых меньшинств и каст⁷³, а также этнических групп, подвергающихся жестокому преследованиям⁷⁴; равно как и злоупотребления в отношении беженцев, мигрантов и просителей убежища⁷⁵. В то же время сообщалось, что платформы пресекают общественную активность лесбиянок, геев, бисексуалов, трансгендеров и квиоров⁷⁶, выступления против репрессивных

⁶³ Monika Bickert, “Hard questions: how we counter terrorism”, 15 June 2017.

⁶⁴ См., например, Twitter Trust and Safety Council и YouTube Trusted Flagger Program.

⁶⁵ Sarah Roberts, *Content Moderation* (University of California at Los Angeles, 2017). См. также материал, представленный организацией «Статья 19», стр. 2.

⁶⁶ Twitter rules and policies (violent extremist groups).

⁶⁷ Facebook community standards (dangerous organizations).

⁶⁸ YouTube policies (violent or graphic content policies).

⁶⁹ См. A/HRC/31/65, пункт 39.

⁷⁰ Facebook community standards (hate speech); Twitter rules and policies (hateful conduct policy).

⁷¹ Amnesty International, *Toxic Twitter: A Toxic Place for Women*; Association for Progressive Communications submission, p. 2; материалы представлены Ассоциацией поддержки прогрессивных коммуникаций, стр. 2.

⁷² Материалы представлены организацией «7amleh» и Ассоциацией поддержки прогрессивных коммуникаций, стр.15.

⁷³ Ijeoma Oluo, “Facebook’s complicity in the silencing of black women”, Medium, 2 August 2017; материалы представлены Центром управления средствами связи, стр. 5 и Ассоциацией поддержки прогрессивных коммуникаций, стр. 11–12.

⁷⁴ Заявление (Statement by) Специального представителя по вопросу о положении прав человека в Мьянме Янги Ли на тридцать седьмой сессии Совета по правам человека, 12 марта 2018 года.

⁷⁵ Материал представлен Ассоциацией поддержки прогрессивных коммуникаций, стр. 12.

⁷⁶ Материал представлен Фондом электронных рубежей, стр. 5.

правительств⁷⁷, сообщения об этнических чистках⁷⁸ и критические высказывания по поводу проявлений расизма и в адрес властных структур⁷⁹.

28. Масштабы и трудности борьбы с ненавистническими высказываниями ставят перед компаниями задачи, требующие долгосрочных усилий, и могут вынуждать их ограничивать такие проявления даже в случаях, когда они явно не связаны с негативным воздействием (поскольку в статье 20 Международного пакта о гражданских и политических правах выступление в пользу ненависти увязывается с подстрекательством). Однако компании обязаны представлять обоснования для таких ограничений, а также доказывать необходимость и соразмерность любых действий в отношении контента (например, удаления контента и приостановления аккаунта). Эффективное и последовательное обеспечение прозрачности политики в отношении ненавистнических высказываний на основе обстоятельного информирования о конкретных случаях способствует выходу на такой уровень понимания, который не может быть достигнут даже с помощью самых подробных пояснений⁸⁰.

29. *Контекст.* Компании обращают особое внимание на важность контекста при оценке применимости общих ограничений⁸¹. Тем не менее такое внимание к контексту не исключает удаления изображений обнаженного тела, имеющих историческое, культурное или образовательное значение⁸²; исторических или документальных свидетельств конфликтов⁸³; доказательств военных преступлений⁸⁴; выступлений против разжигающих ненависть групп⁸⁵; или усилий по оспариванию или опровержению расистских, гомофобных или ксенофобных высказываний⁸⁶. Обстоятельному изучению контекста может препятствовать нехватка у модераторов времени или ресурсов, чрезмерная зависимость от средств автоматизации или недостаточное понимание лингвистических или культурных нюансов⁸⁷. Компании настоятельно призывают пользователей дополнять дискуссионный контент контекстуальными подробностями, однако вопрос об осуществимости и эффективности таких рекомендаций остается открытым⁸⁸.

30. *Требования указания подлинного имени.* В целях борьбы со злоупотреблениями в Интернете некоторые компании вводят требование об указании «подлинного имени»⁸⁹; в то время как другие более гибко подходят к вопросам идентификации⁹⁰. Эффективность требований об указании подлинного имени в качестве гарантии от злоупотреблений в Интернете вызывает сомнения⁹¹. По существу, в результате

⁷⁷ Там же: материалы представлены Ассоциацией поддержки прогрессивных коммуникаций и организацией «7amleh».

⁷⁸ Betsy Woodruff, “Facebook silences Rohingya reports of ethnic cleansing”, The Daily Beast, 18 September 2017; материал представлен организацией «Статья 19», стр. 9.

⁷⁹ Julia Angwin and Hannes Grasseger, “Facebook’s secret censorship rules protect white men from hate speech but not black children”, ProPublica, 28 June 2017.

⁸⁰ См. пункты 52 и 62 ниже.

⁸¹ Twitter, “Our approach to policy development and enforcement philosophy”; YouTube policies (the importance of context); Richard Allan, “Hard questions: who should decide what is hate speech in an online global community?”, Facebook Newsroom, 27 June 2017.

⁸² Материалы представлены «ОБЗЕРВАКОМ», стр. 11 и организацией «Статья 19», стр. 6.

⁸³ Материал представлен организацией «СВИДЕТЕЛЬ», стр. 6–7.

⁸⁴ Там же.

⁸⁵ Материалы представлены Фондом электронных рубежей, стр. 5.

⁸⁶ Материал представлен Ассоциацией поддержки прогрессивных коммуникаций, стр. 14.

⁸⁷ См. Allan, “Hard questions”.

⁸⁸ YouTube policies (the importance of context); Facebook community standards (hate speech).

⁸⁹ Стандарты сообщества «Фейсбук» (использование подлинных личных данных). Следует учитывать, что в зависимости от конкретного случая «Фейсбук» допускает исключения из своего требования об указании подлинного имени, однако такая практика подверглась критике как недостаточно эффективная; материал представлен организацией «Экссес нау», стр. 12. «Байду» даже требует применения идентифицирующей личность данных: пользовательское соглашение «Байду».

⁹⁰ Twitter Help Center, “Help with username registration”; Instagram, “Getting started on Instagram”.

⁹¹ J. Nathan Matias, “The real name fallacy”, Coral Project, 3 January 2017.

применения жесткого требования об указании подлинных имен были раскрыты личности блоггеров и активистов, которые пользовались псевдонимами для собственной защиты, что подвергает их серьезной опасности физической расправы⁹². Блокируются также аккаунты пользователей и активистов из числа лесбиянок, геев, бисексуалов, трансгендеров и квиров, трансвеститов из шоу-бизнеса и пользователей с неанглийскими или необычными именами⁹³. Поскольку анонимность в Интернете зачастую необходима для обеспечения физической безопасности уязвимых пользователей, принципы соблюдения прав человека, являющиеся неотъемлемым компонентом сохранения анонимности, могут подвергаться лишь тем ограничениям, которые обеспечивают защиту личности таких пользователей⁹⁴. Более соразмерным средством защиты личности, прав и репутации других пользователей могут служить узко сформулированные правила обезличивания, ограничивающие способность пользователей ложным или вводящим в заблуждение образом выдавать себя за какое-либо другое лицо⁹⁵.

31. *Дезинформация.* Дезинформация и пропаганда затрудняют доступ к информации и подрывают общее доверие общественности к средствам массовой информации и государственным институтам. Компании ощущают все более острую потребность в борьбе с распространением дезинформации посредством ссылок на фиктивные новостные статьи или веб-сайты третьих сторон, фейковые аккаунты, лживые рекламные объявления, а также путем манипуляции поисковым ранжированием⁹⁶. Однако, поскольку такие прямолинейные действия, как блокировка веб-сайтов или их адресное удаление, могут серьезно ущемлять свободу выражения мнений, компании должны тщательно продумывать любые шаги по борьбе с дезинформацией⁹⁷. Компании применяют широкий круг ответных мер, включая договоренности со сторонними структурами, занимающимися проверкой достоверности сведений, проведение строгой политики в отношении рекламных объявлений, пристальный мониторинг подозрительных аккаунтов, изменение методов курирования контента и алгоритмов поискового ранжирования или обучение пользователей способам распознавания фейковой информации⁹⁸. Некоторые меры, особенно усугубляющие ограничения новостного контента, могут создавать угрозу для независимых и альтернативных новостных источников или сатирического контента⁹⁹. Позиции государственных органов могут отражать чрезмерные ожидания в отношении способности технологий самостоятельно решать такие проблемы¹⁰⁰.

Применяемые компаниями процессы и инструменты модерации

32. *Автоматическое распознавание, удаление и предварительная фильтрация контента.* Под влиянием огромного объема пользовательского контента крупнейшие компании стали разрабатывать инструменты автоматической модерации. Средства автоматизации применяются главным образом для распознавания контента, требующего просмотра людьми, а иногда и для его удаления. Применение автоматизированных средств сканирования музыкального и визуального контента с целью выявления нарушений авторских прав на этапе его загрузки вызывают обеспокоенность с точки зрения возможности чрезмерного блокирования контента, а призывы к использованию фильтрации при загрузке контента террористической и

⁹² Материал представлен организацией «Экссес нау», стр. 11.

⁹³ Dia Kayyali, “Facebook’s name policy strikes again, this time at Native Americans”, Electronic Frontier Foundation, 13 February 2015.

⁹⁴ См. A/HRC/29/32, пункт 9.

⁹⁵ Правила и политика «Твиттера» (политика обезличивания).

⁹⁶ Там же; Allen Babajanian and Christine Wendel, “#FakeNews: innocuous or intolerable?”, Wilton Park report 1542, April 2017.

⁹⁷ Совместное заявление 2017 года.

⁹⁸ Материалы представлены Ассоциацией поддержки прогрессивных коммуникаций, стр. 4–6 и организацией «Статья 19», стр. 4

⁹⁹ Материал представлен Ассоциацией поддержки прогрессивных коммуникаций, стр. 5.

¹⁰⁰ См. сообщение № OL ITA 1/2018. Cf. European Commission, *A Multi-Dimensional Approach to Disinformation: Final Report of the Independent High-level Group on Fake News and Disinformation* (Luxembourg, 2018).

иной направленности грозят введением всеобъемлющих и несоразмерно жестких режимов предварительной цензуры до опубликования материалов¹⁰¹.

33. Средства автоматизации могут быть полезны для компаний, вынужденных анализировать огромные объемы пользовательского контента с использованием самых разных инструментов от фильтров по ключевым словам и программ распознавания спама до алгоритмов проверки хэш-кодов и обработки естественного языка¹⁰². Алгоритм проверки хэш-кодов широко применяется для распознавания изображений сексуальных надругательств над детьми, однако его использование применительно к «экстремистскому» контенту, который обычно требует анализа контекста, трудно осуществимо при отсутствии четких правил в отношении «экстремизма» или его анализа человеком¹⁰³. Сказанное справедливо и в отношении обработки естественного языка¹⁰⁴.

34. *Распознавание и сигнальная маркировка контента пользователями и доверенными модераторами.* Благодаря предоставляемым пользователями сигнальным меткам другие пользователи могут подавать модераторам жалобы на неприемлемый контент. Такие метки обычно не позволяют проводить детальное обсуждение пределов допустимого контента (например, по какой причине контент, который может быть оскорбительным, в целом лучше не удалять)¹⁰⁵. Их также могут злонамеренно использовать для усиления на платформы давления в целях удаления благожелательного контента по отношению к сексуальным меньшинствам и мусульманам¹⁰⁶. Многие компании составляют специальные регистры «доверенных модераторов», как правило, экспертов, влиятельных пользователей и иногда, предположительно, государственных регуляторов¹⁰⁷. В открытом доступе практически отсутствует информация, разъясняющая процесс отбора узкоспециализированных регуляторов, особенности толкования ими юридических или общественных норм или их влияния на принимаемые компаниями решения.

35. *Оценка человеком.* Автоматизированная обработка контента часто дополняется его проверкой человеком, и крупнейшие операторы социальных сетей формируют большие команды модераторов контента для анализа обозначенного сигнальными метками контента¹⁰⁸. Такой контент может перенаправляться модераторам, которые обычно уполномочены принимать, зачастую в течение нескольких минут, решение о приемлемости контента или его удалении либо допуске к размещению. В случаях, когда трудно определить приемлемость того или иного конкретного контента, модераторы могут передать его выше на рассмотрение группам по контролю контента в штаб-квартире компании. В свою очередь должностные лица компании, как правило, группы по вопросам публичной политики или «доверия и безопасности» во взаимодействии с главным юридическим советником будут принимать решения об удалении контента. В целом и в конкретных случаях компании ограниченно раскрывают информацию о дискуссиях по вопросам удаления контента¹⁰⁹.

¹⁰¹ Сообщается, что в Соединенном Королевстве Великобритании и Северной Ирландии разработана программа автоматического распознавания и удаления террористического контента в точке загрузки. Home Office, “New technology revealed to help fight terrorist content online”, 13 February 2018.

¹⁰² Center for Democracy and Technology, *Mixed Messages? The Limits of Automated Media Content Analysis* (November 2017), p. 9.

¹⁰³ Материал представлен Открытым технологическим институтом, стр. 2.

¹⁰⁴ Center for Democracy and Technology, *Mixed Messages?*, p. 4.

¹⁰⁵ По вопросу применения пользовательских сигнальных меток см. в целом Kate Crawford and Tarleton Gillespie, “What is a flag for? Social media reporting tools and the vocabulary of complaint”, *New Media and Society*, vol. 18, No. 3 (March 2016), pp. 410–428.

¹⁰⁶ *Ibid.*, p. 421.

¹⁰⁷ YouTube Help, YouTube Trusted Flagging Program; YouTube Help, “Get involved with YouTube contributors”.

¹⁰⁸ См. Sarah Roberts, “Commercial content moderation: digital laborers’ dirty work”, *Media Studies Publications*, paper 12 (2016).

¹⁰⁹ Cf. Wikipedia: BOLD, revert, discuss cycle. Reddit moderators are encouraged to offer “helpful rule explanations, tips and links to new and confused users” (Reddit Moddiquette).

36. *Меры в отношении аккаунтов или контента.* Наличие неприемлемого контента может стать причиной принятия компанией целого ряда мер. Компании могут ограничивать удаление контента одной единственной юрисдикцией, несколькими юрисдикциями или всей платформой либо комплексом платформ. Они могут вводить ограничения по возрасту, объявлять предупреждения или принимать меры по демонетизации контента¹¹⁰. Нарушения могут повлечь за собой временное блокирование аккаунта, а повторные нарушения могут стать причиной его деактивации. В очень небольшом числе случаев, не подпадающих под действие законодательства об авторских правах, компании применяют процедуры «встречных уведомлений», которые позволяют пользователям размещать контент в порядке опротестования решений о его удалении.

37. *Уведомление.* Одна из широко распространенных причин жалоб состоит в том, что пользователи, размещающие контент, о котором они уведомили, или лица, жалующиеся на злоупотребления, не могут получить никакого уведомления о его удалении или о других мерах¹¹¹. Даже в тех случаях, когда компании направляют уведомления, они обычно всего лишь ставят в известность о принятых мерах и об их общих основаниях. По крайней мере одна компания пыталась в своих уведомлениях сообщать более конкретную информацию, однако остается неясным, обеспечивает ли включение в массово рассылаемые уведомления дополнительных подробностей достаточное объяснение принятых мер¹¹². Транспарентность и уведомления дополняют друг друга: неукоснительное обеспечение транспарентности на оперативном уровне, способствующее повышению осведомленности пользователей о применяемых платформой подходах к удалению контента, в отдельных случаях уменьшает потребность в направлении уведомлений, а более низкий уровень общей транспарентности повышает вероятность того, что при отсутствии уведомлений, составленных с учетом конкретных обстоятельств, пользователи не смогут понять причины удаления контента.

38. *Обжалования и средства правовой защиты.* Платформы разрешают обжаловать определенный круг мер от удаления профилей и страниц до удаления конкретных постов, фото- или видеоматериалов¹¹³. Однако даже при подаче обжалований пользователи, как представляется, располагают настолько ограниченными или несвоевременными средствами правовой защиты, что они фактически у них отсутствуют или в любом случае остаются непонятными для большинства пользователей и даже для экспертов, представляющих гражданское общество. Так, например, восстановление контента может оказаться недостаточной реакцией в случае, если его удаление причинило разместившему его лицу прямой репутационный, физический, моральный или финансовый вред. Подобным образом приостановление действия аккаунтов или удаление контента во время общественных протестных выступлений или диспутов может существенно повлиять на политические права и при этом не сопровождаться предоставлением компаниями средств правовой защиты.

Транспарентность

39. В порядке обеспечения транспарентности компании готовят отчеты, содержащие обобщенные сведения о поступивших от правительства требованиях об удалении контента и предоставлении пользовательских данных. Такие отчеты свидетельствуют об оказываемом на компании давлении. Составление отчетов в порядке обеспечения транспарентности позволяет определить по каждой стране

¹¹⁰ YouTube policies (nudity and sexual content policies); YouTube Help, “Creator influence on YouTube”.

¹¹¹ Материалы представлены организацией «Статья 19», стр. 7 и Ассоциацией поддержки прогрессивных коммуникаций, стр. 16.

¹¹² См. <https://twitter.com/TwitterSafety/status/971882517698510848/>.

¹¹³ Electronic Frontier Foundation and Visualizing Impact, “How to appeal”, onlinecensorship.org. Facebook and Instagram allow only the appeal of account suspensions. Материал представлен веб-сервисом «ГитХаб», стр. 6.

количество законных требований об удалении контента¹¹⁴, число запросов, по которым были приняты определенные меры или ограничен контент¹¹⁵, и, все чаще, содержание или примеры отдельных правовых баз¹¹⁶.

40. Однако согласно выводам одного из ведущих обзоров прозрачности в Интернете компании раскрывают «самый минимальный объем информации о методах разработки и применения частных правил и механизмов самостоятельного и совместного регулирования»¹¹⁷. В частности, «невероятно скудно» предоставляется информация о принятых мерах во исполнение частных просьб об удалении контента в соответствии с условиями предоставления услуг¹¹⁸. Стандарты контента формулируются весьма широко, сохраняя за компаниями свободу принятия решений, которые они недостаточно разъясняют. В результате контроля со стороны социальных сетей и общественности компании были вынуждены дополнять свои общие программные установки размещением в блогах разъяснительных постов¹¹⁹ и отдельными гипотетическими примерами¹²⁰, однако они не обеспечивают объяснение нюансов разработки и применения внутренних правил¹²¹. Если с условиями предоставления услуг в целом можно ознакомиться на местных языках, то отчеты о прозрачности, блоги компаний и связанный с ними контент не доступны на местных языках, что еще более усложняет их понимание пользователями, не владеющими английским языком. В связи с этим пользователи, органы государственной власти и гражданское общество часто выражают неудовлетворение по поводу непредсказуемости условий предоставления услуг¹²². Отсутствие достаточного сотрудничества в сочетании с нарастанием критики со стороны общественности вынуждает компании постоянно анализировать свои правила, пересматривать и отстаивать их.

IV. Правозащитные принципы в контексте модерации контента силами компаний

41. Основатель «Фейсбука» недавно выразил надежду на осуществление процесса, в рамках которого компания «сможет более четко отражать ценности проживающих в различных районах общин»¹²³. Такой процесс и соответствующие стандарты заложены в праве прав человека. Частные нормы, разнящиеся в зависимости от бизнес-модели кампании и расплывчатых концепций общественных интересов, создают для пользователей неустойчивые, непредсказуемые и небезопасные условия и способствуют усилению правительственного надзора. Национальные законы неприемлемы для компаний, которые стремятся применять единые нормы для своих пользовательских баз, отличающихся по географическому расположению и культурной среде. Однако в случае их прозрачного и последовательного применения при конструктивном участии пользователей и гражданского общества

¹¹⁴ Twitter Transparency Report: Removal Requests (January–June 2017); Google Transparency Report: Government Requests to Remove Content; 2016 Reddit Inc., Transparency Report. Facebook does not provide the total number of requests received per country.

¹¹⁵ См., например, Facebook Transparency Report (France) (January–June 2017); Google Transparency Report: Government Requests to Remove Content (India); Twitter Transparency Report (Turkey).

¹¹⁶ Ibid.

¹¹⁷ Материал представлен Организацией по классификации цифровых прав, стр. 4. Курсив оригинала.

¹¹⁸ Там же, стр. 10.

¹¹⁹ См. Elliot Schrage, “Introducing hard questions”, Facebook Newsroom, 15 June 2017; Twitter Safety, “Enforcing new rules to reduce hateful conduct and abusive behavior”, 18 December 2017.

¹²⁰ См., например, YouTube policies (violent or graphic content policies).

¹²¹ Angwin and Grasseger, “Facebook’s secret censorship rules”.

¹²² Материал представлен Организацией по классификации цифровых прав, стр. 10; «ОБЗЕРВАКОМ», стр. 10; Ассоциацией поддержки прогрессивных коммуникаций, стр. 17; Международной федерацией библиотечных ассоциаций и учреждений, стр. 4–5, организацией «Экссес нау», стр. 17; и ЕДРи, стр. 5.

¹²³ Kara Swisher and Kurt Wagner, “Here’s the transcript of Recode’s interview with Facebook CEO Mark Zuckerberg about the Cambridge Analytica controversy and more”, Recode, 22 March 2018.

стандарты прав человека обеспечивают основу для подотчетности государств и компаний перед пользователями вне зависимости от национальных границ.

42. Правозащитная основа предоставляет возможность принимать действенные нормативно-правовые меры в ответ на вводимые государством необоснованные ограничения, при том условии что компании играют по тем же правилам. Руководящие принципы и дополняющий их свод «мягкого права» позволяют компаниям ориентироваться в вопросах о том, какие меры им следует принимать для предотвращения или смягчения чрезмерных правительственных требований об удалении контента. Однако вместе с тем они устанавливают принципы должной осмотрительности, транспарентности, подотчетности и устранения нарушений, которые ограничивают вмешательство интернет-платформ в осуществление прав человека в результате разработки продуктов и политики. Компании, приверженные соблюдению стандартов прав человека во всей своей деятельности, а не только когда они согласуются с их интересами, будут опираться на более прочную основу в своих попытках требовать от государства выполнения этих же стандартов. Кроме того, когда компании более тесно согласуют свои условия предоставления услуг с правом прав человека, государствам труднее использовать их для цензуры контента.

43. Кроме того, правозащитные принципы позволяют компаниям формировать инклюзивную среду, удовлетворяющую различные потребности и интересы своих пользователей, и устанавливать при этом предсказуемые и логичные базовые нормы поведения. На фоне расширяющихся дискуссий по вопросу о том, осуществляют ли компании комплекс посреднических и редакторских функций, право прав человека дает пользователям надежду на то, что они могут опираться на основополагающие нормы для защиты своих мнений сверх устанавливаемых национальным законодательством ограничений¹²⁴. Вместе с тем право прав человека не является настолько негибким или догматичным, что вынуждает компании разрешать выражение мнений, которые могут подрвать права других лиц или способность государств защищать законные интересы национальной безопасности или общественного порядка. В отличие от многочисленных отрицательных факторов, которые могут оказывать более акцентированное воздействие в цифровом пространстве нежели в реальной жизни, включая женоненавистнические или гомофобные преследования, имеющие своей целью заставить замолчать женщин или сексуальные меньшинства, или разжигание всякого рода ненависти, право прав человека не лишает компании их инструментария. Напротив, оно обеспечивает всемирно признанные рамки для разработки такого инструментария и общего лексикона в интересах разъяснения пользователям и государствам их характера, целей и методов применения.

A. Основные стандарты модерации контента

44. Цифровая эпоха позволяет в огромных масштабах быстро распространять информацию, но при этом лишена особенностей жизнедеятельности человека. Согласно Руководящим принципам при оценке необходимости и соразмерности ограничений контента компании могут исходить из размеров, структуры и характерных особенностей создаваемых ими интернет-платформ.

45. *Права человека по умолчанию.* При разработке условий предоставления услуг следует отказываться от дискреционного подхода, основанного на общих и своекорыстных потребностях «сообщества». Вместо этого компаниям следует принять на высоком уровне программные обязательства предоставлять пользователям платформы для формирования мнений, свободного выражения своих убеждений и доступа ко всем видам информации в соответствии с правом прав человека¹²⁵. Такие обязательства должны составлять основу их подходов к модерации контента и

¹²⁴ Материал представлен организацией «Глобальные цифровые партнеры», стр.3; Руководящие принципы, принцип 11.

¹²⁵ Руководящие принципы, принцип 16.

решению таких сложных проблем, как моделируемая пропаганда¹²⁶, а также сбор и обработка пользовательских данных. Компании должны напрямую включать в свои условия предоставления услуг и в «стандарты сообщества» соответствующие принципы права прав человека, обеспечивающие применение мер в отношении контента на основе тех же критериев законности, необходимости и легитимности, которыми обязаны руководствоваться государства при регулировании свободы выражения мнений¹²⁷.

46. *«Законность»*. Правила компаний обычно страдают недостатком четкости и конкретности, которые позволяют пользователям с разумной степенью определенности предсказывать, размещение какого контента выводит их за грань дозволенного. Это особо ярко проявляется в контексте «экстремизма» и «ненавистнических» высказываний, т. е. в тех областях, в которых ограничительные меры легко могут привести к чрезмерному удалению контента при отсутствии тщательной оценки контекста непосредственно людьми. Еще более усложняет понимание общественностью правил размещения контента складывающееся общее представление о необходимости сделать исключение для «событийной ценности» контента¹²⁸. Приветствуя признание общественных интересов, компании вместе с тем должны пояснять, какие факторы подвергаются анализу при определении общественных интересов и какие факторы, помимо общественных интересов, учитываются при расчете событийной ценности. Компании должны дополнять свои усилия по более подробному разъяснению своих правил сводными данными, иллюстрирующими тенденции в области обеспечения выполнения правил, и примерами соответствующих реальных случаев или широким и подробным описанием возможных ситуаций, которые поясняют нюансы толкования и применения конкретных правил.

47. *Необходимость и соразмерность*. Компаниям следует не только более подробно описывать спорные и контекстуальные правила. Они также должны приводить данные и примеры, позволяющие понять факторы, анализируемые при установлении нарушения, степень его тяжести и принятые в ответ на него меры. В контексте ненавистнических высказываний разъяснение методов урегулирования конкретных дел может помочь пользователям лучше понять подход компаний к случаям, в которых сложно провести различие между оскорбительным контентом и разжиганием ненависти, или методику анализа в онлайн-контексте таких соображений, как намерение выступающего с заявлением лица или вероятность насилия. Подробные сведения о принятых мерах будут также создавать основы для оценки степени продуманности вводимых компаниями ограничений. При этом следует разъяснять обстоятельства, при которых они применяют менее строгие ограничения (например, предупреждения, ограничения по возрасту или демонетизацию контента).

48. *Недискриминация*. В интересах обеспечения действенных гарантий недискриминации компаниям следует отходить от формалистических подходов, в рамках которых все защищаемые признаки считаются равно уязвимыми по отношению к злоупотреблениям, притеснениям или иным формам цензуры¹²⁹. По существу, такие подходы представляются несовместимыми с их собственным акцентом на значимость контекста. Вместо этого при разработке или изменении политики или продуктов компании должны активно изучать и учитывать обеспокоенности общин, которые исторически находятся под угрозой цензуры и дискриминации.

¹²⁶ См. Samuel Wooley and Philip Howard, *Computational Propaganda Worldwide: Executive Summary* (Computational Propaganda Research Project working paper No. 2017.11 (Oxford, 2017).

¹²⁷ Материал представлен организацией «Глобальные цифровые партнеры», стр. 10–13.

¹²⁸ См. Joel Kaplan, “Input from community and partners on our community standards”, Facebook Newsroom, 21 October 2016; Twitter rules and policies.

¹²⁹ См., например, Международная конвенция о ликвидации всех форм расовой дискриминации, пункт 4 статьи 1 и пункт 2 статьи 2.

В. Процессы модерации в компаниях и связанная с ними деятельность

Действия в ответ на требования правительств

49. Как явствует из отчетов компаний по вопросам транспарентности, правительства оказывают на них давление в целях принуждения к удалению контента, приостановлению действия аккаунтов, а также к распознаванию и раскрытию информации об аккаунтах. В тех случаях, когда этого требует местное законодательство, у компаний, как выясняется, нет другого выхода, кроме как подчиниться. Однако компании способны разрабатывать инструменты для предупреждения или смягчения рисков для прав человека, создаваемых национальными законами или требованиями, идущими вразрез с международными нормами.

50. *Предупреждение и смягчение.* Компании часто заявляют о своем серьезном отношении к правам человека. Однако дело не может ограничиваться принятием на себя внутренних обязательств и предоставлением общественности временных гарантий при возникновении спорных ситуаций. Компаниям следует также на уровне высшего руководства принимать и обнародовать конкретные направления политики, «которой должны следовать все ее подразделения, включая местные филиалы, при разрешении любых юридически противоречивых вопросов в пользу соблюдения свободы выражения мнений, неприкосновенности частной жизни и других прав человека». Такие обязательства должны быть положены в основу политики и процедур, касающихся толкования или выполнения требований правительств, в интересах обеспечения «наименее жестких ограничений контента»¹³⁰. Компаниям следует добиваться надлежащего представления действующим правительственным органом требований в письменной форме с указанием конкретных существующих нормативных положений об ограничениях¹³¹.

51. При получении противоречивых требований компании должны добиваться их разъяснения или изменения; обращаться за помощью к гражданскому обществу, аналогичным компаниям, соответствующим правительственным органам, международным и региональным организациям и к прочим заинтересованным сторонам; а также изучать все правовые варианты оспаривания полученных требований¹³². В случае же направления государствами в компании требований согласно их условиям предоставления услуг или через посредство иных неправовых механизмов они должны применять к ним процедуры проверки на предмет их соответствия закону и оценивать правомерность таких требований с точки зрения местных законов и стандартов в области прав человека.

52. *Транспарентность.* Перед лицом цензуры и связанных с ней рисков для прав человека пользователи могут принимать осознанные решения о целесообразности и способах выражения своих позиций в социальных сетях только при наличии действительно транспарентных отношений между компаниями и государствами. И в этой связи ощущается потребность в разработке передовых методов обеспечения такой транспарентности. Представляемые компаниями отчеты о требованиях со стороны государства должны дополняться подробными сведениями о типах полученных требований (например, в связи с клеветой, ненавистническими заявлениями, террористическим контентом) и принятых мерах (например, частичное или полное удаление контента, удаление контента в отдельных странах или в глобальном масштабе, приостановление действия аккаунта, удаление в соответствии с условиями предоставления услуг). Помимо этого, компании также должны как можно чаще представлять конкретные примеры¹³³. В составляемые компаниями в порядке обеспечения транспарентности отчеты необходимо включать информацию о

¹³⁰ См. A/HRC/35/22, пункты 66–67.

¹³¹ Материалы представлены Глобальной сетевой инициативой, стр. 3–4 и веб-сервисом «ГитХаб», стр. 3–5.

¹³² См. A/HRC/35/22, пункт 68.

¹³³ См., например, Twitter Transparency Report: Removal Requests (January–June 2017).

требованиях, направляемых правительствами в соответствии с условиями предоставления услуг компаний¹³⁴, а также отражать в них государственно-частные инициативы по ограничению контента, к каковым, например, относятся Кодекс поведения Европейского союза по противодействию незаконным ненавистническим высказываниям в Интернете, правительственные инициативы, в частности создание в Интернете справочных площадок и заключение двусторонних договоренностей по типу достигнутых между «Ютубом» и Пакистаном, а также «Фейсбуком» и Израилем. Компании должны сохранять архивные записи требований, направляемых им по линии таких инициатив, а также переписку между компанией и запрашивающей стороной, и рассматривать возможности достижения договоренностей о передаче копий таких требований на архивное хранение какой-либо третьей стороне.

Разработка правил и продуктов

53. *Должная осмотрительность.* При оценке своих ответных мер на вводимые государством ограничения некоторые компании придерживаются принципа должной осмотрительности в отношении прав человека, однако остается неясным, применяют ли они аналогичные гарантии и для предотвращения или смягчения рисков для свободы выражения мнений, возникающих в результате разработки и реализации их собственной политики¹³⁵. Компании должны разрабатывать четкие и конкретные критерии распознавания видов деятельности, которые дают повод для таких оценок. Помимо анализа политики и процессов модерации контента необходимо производить оценку курирования поступающей от пользователей информации и других форм доставки контента, внедрения новых функций и услуг или изменения существующих, разработки автоматизированных технологий и решений о выходе на рынок, как например договоренностей о разработке специальных версий интернет-платформы для конкретных стран¹³⁶. В представленных ранее отчетах также уточняются вопросы, подлежащие изучению в рамках такой оценки, а также внутренние процессы и потребности в подготовке кадров в интересах учета подобных оценок и сделанных по их итогам выводов в контексте соответствующих операций. Кроме того, такие оценки должны проводиться на текущей основе и корректироваться с учетом изменения обстоятельств или условий деятельности¹³⁷. Подобно Глобальной сетевой инициативе многосторонние инициативы предоставляют компаниям возможность для разработки и совершенствования оценок и других процессов осуществления должной осмотрительности.

54. *Вклад и участие общественности.* Участники консультаций неизменно выражают озабоченность по поводу того, что компании должным образом не взаимодействуют с пользователями и гражданским обществом, особенно в странах глобального Юга. Участие затрагиваемых правообладателей (или их представителей), а также соответствующих местных и профильных экспертов наряду с внутренними процессами принятия решений с должным учетом полученной по обратной связи информации являются неотъемлемыми компонентами проявления должной осмотрительности¹³⁸. Консультации, особенно в таких широких форматах, как обращения к общественности с целью получения от нее замечаний, позволяют компаниям с различных точек зрения анализировать воздействие их деятельности на права человека и побуждают их уделять пристальное внимание вопросу о том, каким образом, казалось бы, безвредные или якобы «учитывающие интересы общин»

¹³⁴ «Твиттер» начал публиковать данные о «незаконных требованиях, которые известные правительственные представители предъявляют к контенту и которые могут нарушать «Правила "Твиттера"» о запрете оскорбительного поведения, поощрения терроризма и нарушений интеллектуальной собственности. Там же. См. также Microsoft, Content Removal Requests Report (January–June 2017).

¹³⁵ Материал представлен организацией «Классификация цифровых прав», стр. 12; Руководящие принципы, принцип 17.

¹³⁶ См. A/HRC/35/22, пункт 53.

¹³⁷ Там же, пункты 54–58.

¹³⁸ См. Руководящие принципы, принцип 18 и A/HRC/35/22, пункт 57.

правила могут оказывать на них существенное, «гиперлокальное» воздействие¹³⁹. Так, например, взаимодействие с проживающими в различных географических районах группами коренного населения может помочь компаниям разработать более эффективные показатели учета культурного и художественно-эстетического контекста при оценке контента, содержащего изображения обнаженного тела.

55. *Транспарентность разработки правил.* Компании, по-видимому, слишком часто внедряют модификации продуктов и правил без обеспечения должной осмотрительности в отношении прав человека или оценки их влияния в конкретных случаях. Они должны как минимум добиваться получения замечаний об оценках их воздействия от заинтересованных пользователей и экспертов в условиях, которые, при необходимости, могут гарантировать конфиденциальность таких оценок. Им также следует недвусмысленно информировать общественность о правилах и процессах их выработки.

Обеспечение выполнения правил

56. *Автоматизированная оценка и оценка человеком.* Автоматизированная модерация контента, введенная под влиянием беспрецедентного увеличения масштабов и объема пользовательского контента, создает очевидные риски принятия в отношении контента мер в нарушение права прав человека. В рамках выполнения своих обязанностей по предупреждению и смягчению воздействия на права человека компании должны учитывать существенные недостатки автоматизации, в частности трудности, возникающие при работе с контекстом, широким диапазоном языковых ключей и значений, а также с лингвистическими и культурными особенностями. Средства автоматизации, разработанные на основе понимания проблем, сформировавшегося в стране пребывания компании, могут подвергаться серьезной дискриминации в масштабах глобальных пользовательских баз. Создаваемые для решения вопросов масштабирования технологии должны как минимум подвергаться строгой проверке и разрабатываться при активном участии пользователей и гражданского общества.

57. Ответственность за формирование надежной и контекстуально-корректной практики модерации контента, предполагающей соблюдение свободы выражения мнений, требует также от компаний укрепления и обеспечения профессионализации в контексте анализа обозначенного сигнальными метками контента силами своих сотрудников. Такие усилия должны включать меры защиты специалистов по модерации контента в соответствии с правозащитными нормами, применимыми к трудовым правам, а также твердую приверженность компаний принципу задействования культурного, лингвистического и прочего экспертного потенциала на каждом рынке, на котором они работают. Кадровый состав руководящих работников и групп по планированию политики компании также следует диверсифицировать, что позволит привлекать местных или профильных экспертов к решению возникающих в связи с контентом вопросов.

58. *Уведомление и обжалование.* Пользователи и эксперты из гражданского общества обычно выражают обеспокоенность по поводу наличия ограниченной информации у лиц, пострадавших от удаления контента или приостановления либо блокировки аккаунтов, а также у лиц, информирующих о таких злоупотреблениях, как женоненавистнические домогательства и доксинг. Отсутствие информации создает обстановку засилья каких-то тайных правил в нарушение критериев прозрачности, конкретности и предсказуемости. Это ущемляет способность отдельных лиц оспаривать принимаемые в отношении их контента меры, а также отслеживать процессы рассмотрения их жалоб на такие меры; однако на практике отсутствие эффективных механизмов обжалования решений об удалении контента более выгодно пользователям, которые сигнализируют о нарушениях, нежели тем, которые

¹³⁹ Chinmayi Arun, “Rebalancing regulation of speech: hyper-local content on global web-based platforms”, Berkman Klein Center for Internet and Society Medium Collection, Harvard University, 2018; Pretoria News, “Protest at Google, Facebook ‘bullying’ of bare-breasted maidens”, 14 December 2017.

размещают посты. Некоторые могут утверждать, что если разрешить обжалование каждого решения по контенту, то это обернется высокими временными и финансовыми затратами. Однако компании могли бы сотрудничать друг с другом и с гражданским обществом для нахождения таких масштабируемых решений, как программы омбудсменов, ориентированные на конкретные компании или целую отрасль. К числу наиболее интересных предложений по таким программам относится идея создания независимого «совета социальных сетей» по образцу советов прессы, которые создают отраслевые механизмы рассмотрения жалоб, а также занимаются развитием правовой защиты от нарушений¹⁴⁰. Такой механизм мог бы заслушивать жалобы отдельных пользователей, отвечающие определенным критериям, а также собирать мнения общественности по поводу регулярно возникающих проблем с модерацией контента, в частности чрезмерной цензуры в той или иной конкретной области.

59. *Возмещение ущерба.* В Руководящих принципах подчеркивается обязанность возмещать «причиненный ущерб» (принцип 22). Однако почти не существует компаний, которые предусматривают меры по возмещению ущерба. Компании должны осуществлять эффективные программы возмещения ущерба, начиная с восстановления в правах и признания ответственности и до внесудебного урегулирования споров в связи с причинением репутационного или иного вреда. Отмечается определенное совпадение правил нескольких компаний в отношении контента, что создает возможности для их сотрудничества в рамках предоставления средств правовой защиты по линии какого-либо возможного совета социальных сетей, других программ омбудсменов или разбирательства дела третьей стороной. В случае регулярного отказа в возмещении может потребоваться вмешательство законодательных или судебных органов.

60. *Автономность пользователей.* Компании разрабатывают механизмы, позволяющие пользователям создавать свою собственную онлайн-среду. К ним относятся подавление или блокировка других пользователей или конкретных видов контента. Аналогичным образом платформы часто разрешают пользователям создавать закрытые или частные группы, модераторами в которых выступают сами пользователи. Поскольку контентные правила в закрытых группах должны соответствовать базовым нормам в области прав человека, платформы должны поощрять такие группы единомышленников с учетом их роли в защите свободы выражения мнений, расширении пространства для уязвимых общин и создании возможностей для тестирования спорных или непопулярных идей. Не следует поощрять требования об указании подлинного имени с учетом их последствий для неприкосновенности частной жизни и безопасности уязвимых лиц¹⁴¹.

61. Нарастающая обеспокоенность по поводу проверяемости, актуальности и полезности онлайн-информации порождает непростые вопросы о методах соблюдения компаниями права на доступ к информации. Как минимум они должны публиковать подробные сведения о своих подходах к кураторству. Если компании классифицируют размещаемый в социальных сетях контент по признаку обмена между пользователями, они должны объяснять собранные ими данные о таких обменах и как они влияют на критерии присвоения рейтингов. Компании должны также предоставлять всем пользователям доступные и реальные возможности для отказа от кураторства со стороны интернет-платформ¹⁴².

¹⁴⁰ См. ARTICLE 19, *Self-regulation and 'Hate Speech' on Social Media Platforms* (London, 2018), pp. 20–22.

¹⁴¹ См. пункт 30 выше.

¹⁴² Так, например, «Фейсбук» разрешает пользователям просматривать материалы в своей «Ленте новостей» в обратном хронологическом порядке, однако, предупреждает, что «в конечном итоге» вернется к выставленным по умолчанию кураторским установкам. Facebook Help Centre, “What’s the difference between top stories and most recent stories on News Feed?”.

Транспарентность решений

62. Несмотря на общее повышение транспарентности в отношении требований правительств об удалении контента, информация о принимаемых мерах в рамках условий предоставления услуг в значительной степени остается закрытой. Компании не публикуют сведения об объеме и типах получаемых в соответствии с ними частных запросов, не говоря уже об их выполнении. Компаниям следует выступать с инициативами по разъяснению воздействия средств автоматизации, практики модерации силами специалистов или обозначения контента пользователями с помощью сигнальных меток либо доверенными модераторами на меры, принимаемые в соответствии с условиями предоставления услуг. Хотя отдельные компании начинают предоставлять о таких мерах определенные данные, вся индустрия в целом должна предпринимать шаги в направлении обнародования более подробных сведений о конкретных и типичных случаях, а также о важных сдвигах в толковании и реализации своей политики.

63. Компании применяют «платформенное право», принимая меры в отношении контента, но практически не разглашая о них информации. В идеале компании должны сформировать своего рода прецедентное право, которое бы помогало пользователям, гражданскому обществу и государствам понимать, каким образом компании истолковывают и применяют свои стандарты. Хотя подобная система «прецедентного права» не будет обеспечивать такой отчетности, которую общественность ожидает от судов и административных органов, централизованная база подробных данных по конкретным делам и примерам столь же эффективно способствовала бы разъяснению правил, как и отчеты по отдельным делам¹⁴³. Совет социальных сетей, уполномоченный на рассмотрение жалоб в рамках всего ИКТ-сектора, мог бы стать авторитетным и независимым механизмом обеспечения транспарентности в этой области.

V. Рекомендации

64. Способность отдельных лиц осуществлять свое право на свободу выражения мнений во всем мире определяется воздействием непрозрачных факторов. Такое обстоятельство требует принятия радикальных мер для обеспечения транспарентности, эффективной подотчетности и приверженности цели исправления существующего положения в интересах защиты способностей отдельных лиц использовать интернет-платформы в качестве форумов для свободного выражения мнений, доступа к информации и участия в общественной жизни. В настоящем докладе намечен целый ряд шагов, включая следующие.

Рекомендации государствам

65. Государства должны отменять любые законы, криминализующие или необоснованно ограничивающие свободу выражения мнений в Интернете и за его пределами.

66. В отличие от субъективного силового регулирования, рациональное регулирование должно стать повседневной практикой, направленной на обеспечение транспарентности компаний и устранение допускаемых ими нарушений, что позволит общественности принимать решения о целесообразности и способах участия в интернет-форумах. Государства должны стремиться к ограничению контента исключительно по решению независимого и беспристрастного судебного органа, а также в соответствии с процессуальными гарантиями и критериями законности, необходимости и легитимности. Государства должны воздерживаться от применения в отношении интернет-посредников несоразмерных санкций, будь то большие штрафы или тюремное заключение, учитывая их негативное влияние на свободу выражения мнений.

¹⁴³ См., например, Madeleine Varner and others, “What does Facebook consider hate speech?”, ProPublica, 28 December 2017.

67. Государствам и неправительственным организациям следует воздерживаться от введения законов или правил, которые бы требовали «упреждающего» мониторинга или фильтрации контента, поскольку ни то, ни другое не соответствует праву на неприкосновенность частной жизни и может быть равнозначно предварительной цензуре.

68. Государствам надлежит воздерживаться от внедрения моделей регулирования, при которых арбитрами по вопросам законного выражения мнений выступают не судебные, а государственные органы. Они должны избегать делегирования компаниям функций арбитров при рассмотрении вопросов контента, поскольку такое делегирование по существу закрепляет в ущерб пользователям примат корпоративных решений над ценностями прав человека.

69. В порядке обеспечения транспарентности государства должны публиковать подробные отчеты о направлении посредникам всех связанных с контентом требований, и привлекать общественность к деятельному участию в обсуждении всех вопросов регулирования.

Рекомендации ИКТ-компаниям

70. Компании должны признать, что воплощением права прав человека являются не изменчивые государственные законы или их собственные частные интересы, а именно общепринятый глобальный стандарт обеспечения свободы выражения мнений на их интернет-платформах, и соответственно пересмотреть свои стандарты в отношении контента. Право прав человека предоставляет компаниям инструменты для формулирования и разработки политики и процессов, обеспечивающих соблюдение демократических норм и противодействие авторитарным требованиям. Реализация такого подхода начинается с принятия правозащитных норм и включает строгую оценку воздействия разработки продуктов и политики на права человека, текущий анализ и пересмотр осуществляемых операций, а также конструктивные консультации с общественностью и гражданским обществом. Руководящие принципы по вопросам предпринимательской деятельности в аспекте прав человека наряду с отраслевыми инструктивными материалами, разрабатываемыми гражданским обществом, межправительственными организациями, Глобальной сетевой инициативой и другими заинтересованными сторонами, предусматривают принципиальные подходы, которых должны придерживаться все интернет-компании.

71. Компаниям надлежит встать на путь применения кардинально новых подходов к транспарентности на всех этапах своей деятельности от разработки правил до осуществления и формирования «прецедентного права», определяющего пределы толкования частных правил. Задача обеспечения транспарентности требует более активного взаимодействия с организациями по вопросам цифровых прав и другими соответствующими структурами гражданского общества, а также отхода от тайных договоренностей с государствами о стандартах контента и их применении.

72. С учетом влияния компаний на общественную жизнь они должны вести открытую деятельность и стать подотчетными обществу. Достойным примером для подражания в части установления минимальных уровней плановости, транспарентности и подотчетности в области модерации коммерческого контента служат созданные по всему миру эффективные и соблюдающие права советы средств массовой информации. Подходы, применяемые третьими сторонами из числа неправительственных организаций, если таковые основываются на стандартах прав человека, могут служить механизмами для обжалования решений и обеспечения возмещения, не будучи сопряжены с высокими затратами, которые отпугивают более мелкие структуры или выходящих на рынок новых игроков. Все сегменты сектора ИКТ, осуществляющие модерацию контента или выполняющие контрольные функции, должны считать одной из своих первоочередных задач создание

**механизмов подотчетности в масштабах всей индустрии
(как, например, совета социальных сетей).**
