

Предубеждения в моделях машинного обучения и аналитике больших данных: последствия для международного уголовного и гуманитарного права

Нема Миланиния*

Аннотация

Развитие технологий мобильной связи и социальных сетей привело к тому, что в современном мире люди больше не способны самостоятельно обрабатывать и использовать весь объем постоянно генерируемой и распространяемой информации. Модели машинного обучения (МО)

* В настоящей статье выражена исключительно точка зрения автора, которая может не совпадать с точками зрения Канцелярии Прокурора МУС или Google. Автор хотел бы выразить благодарность Насрине Баргзи, Алексе Кёниг, Мэттью Кроссу, Бет ван Схак, Марии Соль Бекер, Жерому де Амπτин, Никки Ахмади, Гаянэ Хечумян, Юлии Нузбан и членам редакционной коллегии Международного журнала Красного Креста Брюно Демейеру и Эшу Стэнли-Райану за важный и ценный вклад в настоящую работу и комментарии к ней. Автор настоящей статьи не располагает какими-либо первичными знаниями об исследованиях, описанных в ней.

и инструменты анализа больших данных снимают часть нагрузки, выполняя осмысление такой информации и позволяя получить аналитические выводы, сделать которые иначе было бы невозможно. В контексте международного уголовного права и права прав человека МО используется в различных целях, в том числе для выявления массовых захоронений в Мексике, поиска подтверждений уничтожения жилых домов и школ в Дарфуре, обнаружения сфабрикованных видеороликов и подложных данных, прогнозирования результатов судебных слушаний в Европейском суде по правам человека и сбора доказательств совершения военных преступлений в Сирии. Модели МО также все активнее внедряются в системы вооружений государств, что дает возможность совершенствовать системы прицеливания в аспекте проведения различия между гражданскими лицами, дружественными солдатами и неприятельскими комбатантами или даже обосновывать решения, принимаемые относительно вооруженных нападений.

При этом внедрение такой технологии несет в себе значительный риск. Модели МО и анализ больших данных в существенной степени подвержены обычным человеческим предубеждениям. При наличии предубеждений в моделях МО использование последних потенциально может усугубить существующее расовое, политическое или гендерное неравенство, а также создать вводящую в заблуждение и искаженную картину положения дел на местах. В настоящей статье обсуждаются способы, которыми распространенные человеческие предубеждения могут влиять на работу моделей МО и анализ больших данных, а также изучаются потенциальные правовые последствия таких предубеждений с точки зрения международного уголовного права и международного гуманитарного права.

Ключевые слова: машинное обучение, большие данные, международное уголовное право, международное гуманитарное право, предубеждения, Международный уголовный суд.



Введение

В результате распространения технологий мобильной связи и развития социальных сетей, которые являются инструментом создания и распространения информации, сегодня генерируется больше данных, чем когда-либо ранее, и их объем с каждым днем только растет. Для тех, кто занимается документированием и расследованием международных преступлений или нарушений международного гуманитарного права (МГП) и привлечением к ответственности совершивших их лиц, эти данные потенциально являются кладезем доказательств, позволяющих пролить свет на массовые зверские преступления и выявить тех, кто должен понести ответственность за их совершение. Даже при том, что комиссии ООН по расследованию,

наблюдатели за положением в области прав человека и учреждения международного уголовного права (МУП) (такие как Международный уголовный суд (МУС), Специализированный суд по Косово и Международный остаточный механизм для уголовных трибуналов) привыкли работать с большими и разнообразными наборами данных и большим объемом свидетельских показаний и доказательств¹, они лишь недавно начали по-настоящему осваивать источники больших данных — чрезвычайно крупных наборов данных, изучение которых обычно требует вычислительного анализа (в том числе контент социальных сетей и других цифровых средств массовой информации)². Проще говоря, традиционные методы расследования больше не позволяют человеку обработать и проанализировать весь объем доступной информации. Более того, новые наборы данных часто бывают разнородны и неструктурированы (то есть не имеют определенного единого формата³), как, например, текст, аудио и видео, а также требуют предварительной обработки для извлечения смысловых элементов и обеспечения поддержки метаданных⁴.

Модели машинного обучения (МО) — системы, помогающие программному обеспечению выполнять задачи при отсутствии исчерпывающих программных алгоритмов или правил⁵, — и инструменты анализа больших данных позволяют облегчить такую нагрузку, позволяя осмыслять большие данные и получать аналитические выводы, сделать которые иначе было бы невозможно, в том числе подготавливая нити расследования, демонстрируя закономерности и даже устанавливая сетевые связи или иерархию. Так, например, неправительственные организации (НПО) уже используют модели МО для обнаружения детской порнографии в интернете

- 1 К примеру, в рамках рассмотрения дела Ратко Младича Международным уголовным трибуналом по бывшей Югославии (МТБЮ) было допрошено 377 свидетелей, а также в качестве вещественных доказательств в дело было включено более 10 тыс. видео- и аудиозаписей, фотографий, заключений судмедэкспертов и рукописных документов. ICTY, “Case Information Sheet: Ratko Mladic”, 2020, доступно по адресу: <https://bit.ly/39CgOaa> (все ссылки на интернет-ресурсы приводятся по состоянию на январь 2021 г.).
- 2 Определение «больших данных» см.: Council of Europe, *Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data*, 23 January 2017, п. 3, доступно по адресу: <https://bit.ly/34zMcVn> («Термин “большие данные” обычно обозначает чрезвычайно крупные наборы данных, которые могут быть подвергнуты вычислительному анализу для подготовки заключений о моделях, тенденциях и корреляциях в данных»).
- 3 Неструктурированные данные могут создавать люди или машины. Среди примеров неструктурированных данных, созданных людьми, можно назвать текстовые файлы, электронные письма, данные социальных сетей и мобильной связи. К неструктурированным данным, созданным машинами, относятся в т. ч. спутниковые изображения, научные данные, данные цифрового надзора и данные с различных датчиков. См.: UN Secretary-General, *Data Strategy for Action by Everyone, Everywhere (2020–2022)*, 2020, p. 81, доступно по адресу: <https://bit.ly/3iqCdY2>.
- 4 Эти данные противопоставляются традиционным структурированным данным, таким как данные банковских транзакций, которые обычно отличаются высокой степенью организации и имеют удобный для поиска в реляционной базе данных формат. Ibid, p. 81.
- 5 Ibid., p. 80; *Международный Комитет Красного Креста (МККК)*. Человек должен играть ключевую роль в применении искусственного интеллекта и машинного обучения в ходе вооруженных конфликтов, 6 июня 2019 г., с. 1 и 10, в этом номере Журнала.

и оповещения об этом соответствующих надзорных органов⁶. В Мексике модель МО, спроектированная силами местных и международных НПО, применяется для прогнозирования и обнаружения мест массового захоронения⁷. Исследователи Университета Карнеги — Меллона создали систему видеоанализа, основанную на МО и машинном зрении, которая носит название «Система маркировки событий посредством аналитической обработки медиаматериалов» (E-LAMP) и служит для обнаружения объектов, звуков, речи, текста и различных типов событий (убийств, изнасилований или других преступлений) в собраниях видеоматериалов⁸. Модели МО также применяются Бюро Великобритании по борьбе с мошенничеством в особо крупных размерах для выявления защищенных от раскрытия законом материалов среди миллионов документов, содержание которых раскрывается в рамках расследований, а также Управлением инспекции труда Норвегии для прогнозирования высокого уровня риска на рабочих местах, которые этот орган впоследствии инспектирует⁹. Наконец, еще одним примером является запущенная исследователями проекта «Сирийский архив» система VFRAME, которая фиксирует удары с использованием кассетных боеприпасов в Сирии и Йемене¹⁰. Выполнение большинства описанных выше задач заняло бы у людей много лет; модели МО позволяют сократить это время до нескольких дней.

Кроме того, модели МО находят все более широкое применение в рамках вооруженных конфликтов. Министерство обороны США ведет активную работу по внедрению МО в работу специальных подразделений по сбору разведывательных данных. Обученные с помощью МО программы будут изучать видеоматериалы, снятые с беспилотных летательных аппаратов, и автоматически выявлять враждебные действия для определения боевых целей¹¹. Кроме того, ведомство использует модели МО при командовании и координации действий, чтобы отфильтровывать сведения по различным аспектам и компилировать их в единый источник информации, дающий комплексную картину положения дружественных и недружественных сил, а также помогающий принимать решения относительно нападе-

6 Nikola Todorovic and Abhi Chaudhuri, “Using AI to Help Organizations Detect and Report Child Sexual Abuse Material Online”, *The Keyword*, 3 September 2018, доступно по адресу: <https://bit.ly/2HJx9Qi>.

7 Mimi Onuoha, “Machine Learning Is Being Used to Uncover the Mass Graves of Mexico’s Missing”, *Quartz*, 19 April 2017, доступно по адресу: <https://bit.ly/31PxFDo>.

8 Jay D. Aronson, Shicheng Xu and Alex Hauptmann, *Video Analytics for Conflict Monitoring and Human Rights Documentation: Technical Report*, Carnegie Mellon University, July 2015, доступно по адресу: <https://bit.ly/2LXJhiH>.

9 Annette Vestby and Jonas Vestby, “Machine Learning and the Police: Asking the Right Questions”, *Policing: A Journal of Policy and Practice*, 14 June 2019, p. 5, доступно по адресу: <https://bit.ly/3nVyLp8>.

10 Karen Hao, “Human Rights Activists Want to Use AI to Help Prove War Crimes in Court”, *MIT Technology Review*, 25 June 2020, доступно по адресу: <https://bit.ly/3e9M1mX>.

11 Congressional Research Service, *Artificial Intelligence and National Security*, 10 November 2020, p. 10, доступно по адресу: <https://bit.ly/2XNcEH5>.

ния¹². Наконец, МО внедряется в автономные системы оружия, в том числе как средство выбора и захвата целей¹³.

При этом модели МО, как и все прочие инструменты анализа больших данных, по своей природе не являются объективными. Инженеры обучают модели на данных, и участие человека в процессе подбора и подготовки таких данных может приводить к отражению предубеждений в прогнозах, которые дают такие модели¹⁴. Так происходит потому, что на сбор данных часто влияют различные предубеждения, из-за которых некоторые группы или события оказываются недостаточно или избыточно представлены. Это особенно актуально, если речь идет о больших данных, так как многие из наборов данных в таком случае не отличаются высоким качеством организации, обычно присущим данным статистических исследований, а наоборот, являются побочным продуктом других действий, сформированным для других (часто исключительно операционных) целей¹⁵. Например, модель МО для распознавания изображений, созданная одним из специалистов по теории вычислительных машин и систем Виргинского университета, несоразмерно часто связывала изображения кухонь с женщинами¹⁶. Причиной таких ассоциаций было то, что на использованных для обучения ПО фотографиях некоторые виды деятельности, такие как приготовление пищи и уборка, чаще выполнялись женщинами, а не мужчинами. Это пример характерного гендерного предубеждения. Под влиянием подобных предубеждений результаты работы моделей МО или других инструментов анализа больших данных могут значительно искажаться.

По сей день отсутствуют эффективные международные правовые механизмы, которые бы обеспечивали легитимность моделей МО, аналитики больших данных и даже данных социальных сетей с точки зрения МУП и МПП. Судьям Специального трибунала по Ливану в работе пришлось столкнуться с необходимостью сложного анализа телекоммуникаций, и, например, Судебная камера не рассмотрела ни один из особо острых вопросов, касающихся предубеждений в процессах сбора и интерпретации телекоммуникационных данных (более того, ни один из таких вопросов даже не был поднят). Наиболее хронологически близкий ко времени написания настоящей статьи пример — дело Аль-Верфалли, в рамках которого ордер МУС на арест был выдан по большей части на основании сведений, размещенных в Facebook и YouTube, но без использования

12 Theresa Hitchens, “Air Force Expands 5G as It Transforms to Multi-Domain Ops: Donovan”, *Breaking Defense*, 4 September 2019, доступно по адресу: <https://breakingdefense.com/2019/09/air-force-expands-5g-as-it-transforms-to-multi-domain-ops-donovan/>.

13 Michael N. Schmitt, “Autonomous Weapons Systems and International Humanitarian Law: A Reply to the Critics”, *Harvard National Security Journal: Features Online*, 5 February 2013, p. 28, доступно по адресу: <https://bit.ly/3ip5pyh>.

14 Facebook, *Facebook’s Civil Rights Audit — Final Report*, 8 July 2020, p. 76, доступно по адресу: <https://bit.ly/3nVlCwk>.

15 МЖКК (примечание 5 выше), с. 10.

16 Tom Simonite, “Machines Taught by Photos Learn a Sexist View of Women”, *Wired*, 21 August 2017, доступно по адресу: <https://bit.ly/3qvxaIm>.

моделей МО¹⁷. Несмотря на то что активисты назвали это дело важнейшим событием — первым прецедентом выдачи международного ордера на арест, выданным на основании контента социальных сетей¹⁸, решение касалось только ордера на арест, что соответствует низшему уровню в иерархии доказательств по МУП — разумному основанию полагать¹⁹. Ни одно доказательство из социальных сетей, которыми был обоснован ордер на арест Аль-Верфалли, не было проверено в рамках перекрестного допроса свидетелей или на соответствие более высокому уровню доказательств, требуемому для вынесения обвинительного приговора (отсутствие обоснованного сомнения)²⁰. В силу более низкого уровня доказательств, требуемого для вынесения таких предварительных решений, в правоприменительной практике МУП также не учитываются технологические новшества, максимально упрощающие манипулирование данными, в том числе фотографиями и видеозаписями²¹. Не учитывает она и потенциальное воздействие человеческих предубеждений.

Отсутствие эффективных международных правовых механизмов для решения таких вопросов в целом отражает то обстоятельство, что международные правовые учреждения пока не сталкиваются в рассмотрении дел с количеством релевантных случаев, достаточным для того, чтобы формирование такого представления о социальных сетях, больших данных и МО стало очевидно необходимым. Однако данный пробел дает этим учреждениям и их следователям, аналитикам и прокурорам возможность выработать нормы и практические методы, которые бы могли опираться на опыт национальных правоприменительных органов в сфере МО и больших данных. Основной задачей здесь является создание норм, регулирующих потенциальное воздействие человеческих предубеждений. До настоящего времени научное сообщество не исследовало эти предубеждения детально и не обращало на них достаточно внимания. Эта мысль справедлива и в отношении возможного воздействия предубеждений на ход международных уголовных расследований и потенциальных правовых последствий с точки зрения МУП и МГП. Настоящая статья призвана заполнить этот пробел в научном знании. В первой части статьи обобщаются сведения о наиболее распространенных типах человеческих предубеждений, воздействующих на модели МО и анализ больших данных, а также об их возможных последствиях. Во второй части рассматриваются потенциальные

17 ICC, *Prosecutor v. Mahmoud Mustafa Busayf Al-Werfalli*, Case No. ICC-01/11-01/17, Warrant of Arrest (Pre-Trial Chamber I), 15 August 2017.

18 См., например: Emma Irving, “And So It Begins... Social Media Evidence in an ICC Arrest Warrant”, *Opinio Juris*, 17 August 2017, доступно по адресу: <https://bit.ly/3kvEtNI>.

19 Римский статут Международного уголовного суда, док. ООН A/CONF.183/9, 17 июля 1998 г. (вступил в силу 1 июля 2002 г.) (Римский статут), ст. 58(1).

20 Там же, ст. 66(3).

21 Единственным условием является доказательство обвинением вины на разумных основаниях. ICC, *Prosecutor v. Omar Hassan Ahmad Al Bashir*, Case No. ICC-02/05-01/09, Decision on the Prosecution's Application for a Warrant of Arrest against Omar Hassan Ahmad Al Bashir (Pre-Trial Chamber I), 4 March 2009, paras 32–34.

правовые последствия этих предубеждений с точки зрения МГП и МУП. Призмой для анализа служит Римский статут МУС (Римский статут).

Предубеждения, распространенные в моделях машинного обучения и аналитике больших данных

В наборах данных часто содержатся предубеждения, из-за которых могут быть несправедливо ущемлены интересы определенных групп или чрезмерно повышена приоритетность определенных действий в ущерб другим. Модели МО или инструменты анализа больших данных, обученные на таких наборах, могут унаследовать эти предубеждения²². В настоящем разделе рассмотрены человеческие предубеждения, которые чаще других встречаются в наборах данных для моделей МО и, соответственно, с большей вероятностью могут повлиять на исход расследований в области МУП и разбирательств в области МГП: скрытые предубеждения, ошибки выбора, предубеждения в отчетности, ошибки групповой атрибуции и предубеждения в пользу автоматизированных систем²³. Относительно каждого типа предубеждений в статье приводится информация о способах их воздействия на модели МО или анализ больших данных, особенно в контексте международных уголовных расследований или нарушений МГП.

Скрытые предубеждения

Скрытые предубеждения появляются тогда, когда допущения делаются на основании психологических моделей отдельных людей и личного опыта, которые могут быть неприменимы в более широких масштабах. Эти предубеждения часто несут в себе дискриминационную составляющую, например скрытые расовые или гендерные предпочтения. Так, в 2018 году в Amazon обнаружили, что алгоритмы скрининга резюме для отбора кандидатов обучались на данных, содержащих скрытые предубеждения против кандидатов-женщин. Вследствие этого алгоритм выставлял более низкие оценки резюме, содержащим слова с корнем «жен-», например «капитан женского шахматного клуба»²⁴.

22 UN Institute for Disarmament Research, *Algorithmic Bias and the Weaponization of Increasingly Autonomous Technologies*, 2018, p. 3, доступно по адресу: <https://bit.ly/3nPmiTX>.

23 Следующий перечень содержит лишь краткую выборку предубеждений, которые часто обнаруживаются в наборах данных, применяемых для МО. Перечень не является исчерпывающим. В каталоге когнитивных искажений (предубеждений) в Википедии перечислено более 100 различных типов человеческих предубеждений, которые могут влиять на формирование суждений, и соответственно, на модели МО; см. *Википедия*. Список когнитивных искажений, доступно по адресу: https://ru.wikipedia.org/wiki/Список_когнитивных_искажений. См. также: Forensic Science Regulator, *Cognitive Bias Effects Relevant to Forensic Science Investigations*, 4 April 2018, доступно по адресу: <https://bit.ly/3bNOQe9>.

24 Jeffrey Dastin, "Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women", *Reuters*, 10 October 2018, доступно по адресу: <https://reut.rs/2HItB0B>.

Скрытые предубеждения могут принимать различные формы. К распространенным скрытым предубеждениям относится предпочтение информации, подтверждающей сложившуюся точку зрения. Его суть заключается в том, что люди или создатели моделей неосознанно обрабатывают данные так, чтобы они подтверждали ранее существующие взгляды и гипотезы²⁵. В контексте международных расследований это предубеждение может заставить следователей или прокуроров не обратить внимания на потенциально исключающие вину обстоятельства или снизить их значимость. Как результат, сбора и раскрытия таких данных может не произойти²⁶. Похожим образом в тяжелых условиях театра боевых действий комбатанты, находящиеся под влиянием предубеждений, могут принять гражданских лиц или объекты гражданской инфраструктуры за военные цели, как произошло в 1988 году, когда крейсер ВМС США «Винсеннес» по ошибке сбил иранский пассажирский самолет, установив, что поведение воздушного судна схоже с маневрами боевого самолета F-14²⁷.

С предпочтением информации, подтверждающей сложившуюся точку зрения, тесно связаны другие виды скрытых предубеждений — выборочная обработка информации, непоколебимая убежденность и уход от когнитивного диссонанса. Каждое из трех предубеждений может приводить к тому, что прокуроры, следователи, военнослужащие или аналитики будут игнорировать ценную информацию, которая противоречит ранее сложившейся у них позиции по делу. Выборочно обрабатывая информацию, люди переоценивают значимость информации, согласующейся с ранее сформировавшимися у них убеждениями, и недооценивают значимость информации, которая может их опровергнуть²⁸. Непоколебимая убежденность — термин, которым описывают склонность людей придерживаться какой-либо версии или теории даже после того, как обосновывающие ее доказательства были опровергнуты²⁹. Наконец, желая уйти от когнитивного диссонанса, люди могут корректировать свои взгляды с тем, чтобы они соответствовали установившемуся личному восприятию³⁰. Как отмечается, такие предубеждения в рамках уголовного преследования могут существенно повлиять на решения, принимаемые прокурором, в том числе относительно расследования и обвинения, презумпции виновности или невиновности или раскрытия исключающих вину обстоятельств.

25 Sonia K. Katyal, “Private Accountability in the Age of Artificial Intelligence”, *UCLA Law Review*, Vol. 66, No. 1, 2019, p. 79; Kate E. Bloch, “Harnessing Virtual Reality to Prevent Prosecutorial Misconduct”, *Georgetown Journal of Legal Ethics*, Vol. 32, No. 1, 2019, p. 5.

26 Alafair S. Burke, “Improving Prosecutorial Decision Making: Some Lessons of Cognitive Science”, *William & Mary Law Review*, Vol. 47, No. 5, 2006, pp. 1603–1604.

27 Peter Margulies, “The Other Side of Autonomous Weapons: Using Artificial Intelligence to Enhance IHL Compliance”, in Ronald T. P. Alcalá and Eric Talbot Jensen (eds), *The Impact of Emerging Technologies on the Law of Armed Conflict*, Oxford University Press, Oxford, 2019, pp. 147, 158–159.

28 A. S. Burke (примечание 26 выше), pp. 1594, 1596–1599; Alafair S. Burke, “Commentary: Brady’s Brainteaser: The Accidental Prosecutor and Cognitive Bias”, *Case Western Reserve Law Review*, Vol. 57, No. 3, 2007, p. 578.

29 A. S. Burke (примечание 26 выше), pp. 1594, 1599–1601.

30 *Ibid.*, pp. 1594, 1601–1602.

В контексте принятия решений в рамках прокурорской работы источником предубеждения является уверенность в том, что обвиняемый виновен. Как только такая уверенность появляется, предпочтение информации, подтверждающей сложившуюся точку зрения, заставляет прокурора искать информацию, подтверждающую теорию виновности и ставящую под вопрос обстоятельства, которые потенциально могут исключать вину; непоколебимая убежденность же заставляет придерживаться теории виновности даже тогда, когда изначально подтврждавшие ее доказательства более несостоятельны³¹.

Скрытые предубеждения являются весьма острой проблемой в рамках международных уголовных расследований, так как к началу большинства таких расследований новостные агентства, НПО и организации системы ООН успевают подробно осветить те события, которые характеризуются как международные преступления. Так, например, МУС начал расследование преступлений, совершенных против народа рохинья в Мьянме, только в ноябре 2019 года³², несколько лет спустя после первых эпизодов, имевших место в 2016 году, и после публикации организациями системы ООН и различными НПО множества докладов о ситуации с правами человека, в которых задокументированы факты совершения этих преступлений³³. Аналитики МУС полагались на эти доклады, запрашивая разрешение на начало расследования, и следователи Канцелярии Прокурора (КП) продолжают полагаться на них в поисках нитей расследования и формирования версии следствия³⁴. В то же время такие доклады могут и будут воздействовать на профессиональное заключение следователя относительно того, как и кем были совершены преступления. Подобные опасения были совсем недавно высказаны Специальным докладчиком по вопросу о внесудебных казнях, казнях без надлежащего судебного разбирательства или произвольных казнях при расследовании смерти Джамала Хашогги:

Ко времени начала дознания о большинстве убийств было уже сообщено и была установлена вероятная ответственность ряда лиц. Риск предпочтения информации, подтверждающей сложившуюся точку зрения (стремления к укреплению какой-либо гипотезы посредством

31 A. S. Burke (примечание 26 выше), p. 1614.

32 ICC, *Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar*, Case No. ICC-01/19, Decision Pursuant to Article 15 of the Rome Statute on the Authorisation of an Investigation into the Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar (Pre-Trial Chamber III), 14 November 2019.

33 См., например: *Совет по правам человека*. Доклад независимой международной миссии по установлению фактов в Мьянме, док. ООН A/HRC/39/64, 27 августа 2018 г.; Médecins Sans Frontières, "No One Was Left": *Death and Violence against the Rohingya in Rakhine State, Myanmar*, 9 March 2018, доступно по адресу: <https://bit.ly/3edvEFV>.

34 ICC, *Situation in the People's Republic of Bangladesh/Republic of the Union of Myanmar*, Case No. ICC-01/19, Request for Authorisation of an Investigation Pursuant to article 15 (Pre-Trial Chamber III), 4 July 2019.

поиска подтверждающих ее доказательств и при игнорировании других, не сопоставимых с ней доказательств), был особенно высок³⁵.

Описанные выше обстоятельства повысили риск возникновения такого предубеждения, особенно если принять во внимание то, сколь много информации о ситуации было размещено в социальных сетях и на других платформах. Время от времени это предубеждение также распространяется на процесс взаимодействия международного сообщества с МУС. В силу того, что согласно Программному документу КП по предварительным разбирательствам их следует проводить, исходя из информации, полученной КП, в сочетании с материалами из открытых источников³⁶, существует внешнее (и, предположительно, иногда внутреннее) ожидание, что расследование должно ответить на вопросы, поднятые в заключении предварительной экспертизы.

Использование моделей МО, построенных на наборах данных, которые отягощены скрытыми предубеждениями, может приводить к значительным последствиям с точки зрения МГП, например обуславливать принятие людьми ошибочных решений о выборе в качестве боевых целей объектов гражданской инфраструктуры. В 2015 году в ходе нападения США на здание организации «Врачи без границ» (ВБГ) в Кундузе, Афганистан, внимание личного состава как сухопутных войск, так и авиации привлекли «арочные ворота» в структуре здания, а также территория вокруг него, «окруженная стеной по внешнему периметру, за которой находилось несколько построек». Исходя из этих особенностей, военнослужащие пришли к выводу о том, что перед ними действующая база «Талибан»³⁷. Однако перечисленные особенности характерны для архитектуры всех зданий в этом регионе, поэтому их наличие не должно было служить определяющим доводом при выборе боевой цели. К сожалению, военнослужащие, ответственные за принятие решений при прицеливании, подтвердили выбор здания ВБГ в качестве боевой цели ввиду сложившихся у них ранее представлений о том, как выглядят военные объекты, подлежащие атаке.

Как в ходе международных расследований, так и при вынесении судебных определений в области МГП использование моделей МО, основанных на наборах данных со скрытыми предубеждениями, с большой вероятностью приводит к искажению аналитических выводов, которые

35 Human Rights Council, *Annex to the Report of the Special Rapporteur on Extrajudicial, Summary or Arbitrary Executions: Investigation into the Unlawful Death of Mr. Jamal Khashoggi*, UN Doc. A/HRC/41/CRP.1, 19 June 2019, para. 37.

36 ICC, OTP, *Policy Paper on Preliminary Examinations*, November 2013, paras 79, 80, 104, доступно по адресу: <https://bit.ly/3nXQ2y6>. См. также: ICC, *Proposed Programme Budget for 2021 of the International Criminal Court*, ICC-ASP/19/10, 10 September 2020, para. 128, доступно по адресу: <https://bit.ly/2LxHkJZ>.

37 US Central Command, "Summary of the Airstrike on the MSF Trauma Center in Kunduz, Afghanistan on October 3, 2015", 29 April 2016, p. 389; Matthew Rosenberg, "Pentagon Details Chain of Errors in Strike on Afghan Hospital", *New York Times*, 29 April 2016, доступно по адресу: <https://nyti.ms/3irFBBJ>; P. Margulies (примечание 27 выше), pp. 149–150.

могут создавать нереалистичную картину преступных связей, отношений и закономерностей. Как изложено ниже, такие предубеждения могут оказать воздействие на обязательства о раскрытии и обязательство прокурора по установлению истины, а также потенциально усугубить стереотипы, которые обычно присутствуют в контенте, создаваемом пользователями (КСП). Предубеждения также могут оказать влияние на решения при прицеливании и привести к ошибочным нападениям на гражданских лиц и объекты гражданской инфраструктуры. Более того, скрытые предубеждения могут подкреплять предубеждения международного уровня в процедурах прицеливания, что несет угрозу для гражданских лиц; примером может служить внедрение стратегии нападения на террористов с использованием БПЛА в код алгоритма³⁸. Скрытые предубеждения чаще других форм предубеждений встречаются в расследованиях уголовных дел и разбирательствах в области МПП, а также требуют самых серьезных мер урегулирования последствий.

Ошибка отбора

Ошибки отбора происходят в ситуациях, когда наборы данных, используемые для обучения моделей МО или для анализа, отбираются таким образом, что в итоге их реальное распределение не отражается³⁹. Например, стоит задача создать модель, которая должна обеспечивать работу камер видеонаблюдения, но ее обучали только на данных, полученных в ночное время. По причине этого в модель была включена ошибка отбора, которая может отрицательно отразиться на эффективности работы камер в дневное время.

Ошибка отбора может проявляться по-разному и принимать различные формы. Погрешность охвата — одна из форм ошибки отбора, возникающая в случаях, когда опорный набор данных неполон, и выборка данных, на основе которой делаются выводы, не является репрезентативной для целевой популяции⁴⁰. Говоря конкретнее, при такой погрешности популяция больших данных *не* совпадает с целевой популяцией. Например, если аналитики-криминалисты используют модель МО для выявления закономерностей, но не включают в нее данные по преступлениям, совершаемым против детей, модель будет иметь погрешность охвата и, вероятно, не сможет выявлять закономерности применительно к таким преступлениям. Искажение, связанное с отсутствием ответов, или ошибка неполного участия — это разновидность ошибки отбора, которая возникает, когда пользователи из определенных групп отказываются от участия в процессе, поэтому данные в конечном счете становятся нерепрезентативными в силу

38 Ben Tarnoff, “Weaponised AI is Coming. Are Algorithmic Forever Wars Our Future?,” *The Guardian*, 11 October 2018, доступно по адресу: <https://bit.ly/3qz3hqT>.

39 Patrick Ball, “The Bigness of Big Data”, in Philip Alston and Sarah Knuckey (eds), *The Transformation of Human Rights Fact-Finding*, Oxford University Press, Oxford, 2015, pp. 425, 436–437.

40 Joann Stonier, “Fighting AI Bias — Digital Rights Are Human Rights”, *Forbes*, 19 March 2020, доступно по адресу: <https://bit.ly/35FGKzH>.

наличия пробелов в процессе сбора данных⁴¹. Эта форма предубеждений преобладает в тех ситуациях, когда ущемленные в правах или традиционно недостаточно представленные группы не доверяют процессу и, соответственно, с меньшей вероятностью бывают готовы в нем участвовать⁴². Кроме того, такая ошибка часто возникает в рамках внутренних вооруженных конфликтов или в других ситуациях насилия, когда отсутствие доверия к учреждениям и органам власти может серьезно отразиться на уровне участия представителей уязвимых и пострадавших сообществ⁴³.

Ошибка выборки — это разновидность ошибки отбора, которая проявляется, когда сбор данных по целевой группе населения производится не случайным образом, а формируется только на основе выборки из определенного подмножества данных целевой группы⁴⁴. Выводы, подготавливаемые моделью, в результате в большей степени отражают положение именно в той части популяции, по которой сформирована выборка. Наконец, ошибка масштаба события — это разновидность ошибки отбора, которая связана с вероятностью учета «того или иного события в отчетности в связи с его масштабом: крупные события чаще получают огласку, чем менее масштабные»⁴⁵. В этом отношении события, находящие большее отражение в наборе данных, получают больше внимания, чем столь же тяжкие, но менее масштабные. Например, убийства, совершаемые в дневное время и широко освещаемые в социальных сетях, будут в большей степени влиять на модели МО, нежели акты сексуального насилия, на которые общественность обращает меньше внимания или не обращает его вовсе.

Ошибки отбора могут вызывать особенно острые проблемы при расследовании массовых зверств или нарушений прав человека. Как отмечает Джей Аронсон:

В случае нарушений прав человека число жертв всегда относительно мало в сравнении с общей численностью популяции, и они часто так или иначе ущемлены в правах (например, имеют ограниченный доступ к устройствам сетевой связи и интернету в силу бедности или проживания в сельской местности). Данное обстоятельство повышает вероятность того, что произвольная выборка из большого массива данных

41 Hanna Tolonen, Miika Honkala, Jaakko Reinikainen, Tommi Härkänen and Pia Mäkelä, “Adjusting for Non-Response in the Finnish Drinking Habits Survey”, *Scandinavian Journal of Public Health*, Vol. 47, No. 4, 2019, p. 470.

42 Andrea F. de Winter, Albertine J. Oldehinkel, René Veenstra, J. Agnes Brunnekreef, Frank C. Verhulst and Johan Ormel, “Evaluation of Non-Response Bias in Mental Health Determinants and Outcomes in a Large Sample of Pre-Adolescents”, *European Journal of Epidemiology*, Vol. 20, No. 2, 2005.

43 Sam Whitt, “Institutions and Ethnic Trust: Evidence from Bosnia”, *Europe-Asia Studies*, Vol. 62, No. 2, 2010.

44 Andrew D. Selbst, “Disparate Impact in Big Data Policing”, *Georgia Law Review*, Vol. 52, No. 1, 2017, pp. 109, 134–135.

45 Megan Price and Patrick Ball, “Big Data, Selection Bias, and the Statistical Patterns of Mortality in Conflict”, *SAIS Review*, Vol. 36, No. 1, 2014, p. 11.

не будет учитывать как минимум некоторые (если не многие) случаи нарушения⁴⁶.

Таким образом, если заранее не обеспечить репрезентативность набора данных относительно групп населения, которые обычно недостаточно представлены в статистике, существует риск того, что модели МО и другие инструменты анализа больших данных не будут учитывать такие преступления в принципе.

Ошибки отбора также проявляются в тех случаях, когда демографический состав работников, выполняющих анализ и ввод данных, не отличается репрезентативностью. На этот факт ссылались авторы недавнего заключения о проверке Facebook на соблюдение гражданских прав, где отмечалось, что «ключевое условие обеспечения справедливости алгоритмов — заставить компании целенаправленно работать над разнообразием кадрового состава, в том числе той его части, которая работает над алгоритмом Facebook»⁴⁷. Применив это условие к оценкам с точки зрения МУП или МГП, получим следующее: если коллектив исследователей или аналитиков, задействованных при расследовании МУС относительно событий в Афганистане, состоит исключительно из людей, которые близко не знакомы с культурой и языками населения Афганистана, то выводы расследования почти неизбежно будут содержать преувеждения. Группа следователей, состоящая только из англоговорящих людей, не обладающих опытом работы в Афганистане и не понимающих дари или пушту, с большей вероятностью сосредоточится на доказательствах, связанных с ответственностью международных сил, например вооруженного контингента США, — ведь такие доказательства представлены на английском языке и потому более доступны и понятны, — чем на происшествиях, задокументированных исключительно на пушту или дари. Расследования развиваются именно таким образом, даже несмотря на то, что ответственность за гибель большинства жертв из числа гражданских лиц, очевидно, несут внутренние силы Афганистана⁴⁸. Обратная ситуация также не является приемлемой, так как результаты расследования, выполненного группой лиц исключительно афганского происхождения, тоже будут содержать ошибки отбора — только иные. Любая модель МО или инструмент анализа, основанный на сформированном такой группой наборе данных, тоже будет включать преувеждения, отражающие точку зрения следователей.

Ошибки отбора несут в себе следующий риск: если данные не отражают реального распределения событий, то обученная на них модель МО

46 Jay D. Aronson, “Mobile Phones, Social Media, and Big Data in Human Rights Fact-Finding: Possibilities, Challenges, and Limitations”, in P. Alston and S. Knuckey (eds) (примечание 39 выше), pp. 441, 447.

47 Facebook (примечание 14 выше), p. 80.

48 См.: UN Assistance Mission in Afghanistan, *Afghanistan: Protection of Civilians in Armed Conflict, 2019, 2020*, pp. 5–6, доступно по адресу: <https://bit.ly/3e8ObmQ> (где отмечается, что «действия антиправительственных элементов в 2019 г. оставались причиной большинства (62%) жертв среди гражданского населения»).

впитает и укрепит этот тип предубеждений. Например, ошибка отбора может повлиять на то, соответствуют ли отраженные в данных закономерности или масштаб событий реальности. Данная проблема имеет первостепенное значение при формировании ряда правовых заключений с точки зрения МГП, в том числе при определении того, достаточно ли интенсивны, продолжительны и часты нападения, чтобы квалифицировать их как немеждународный вооруженный конфликт⁴⁹, и с точки зрения МУП, в том числе при определении того, выполнены ли критерии преступлений против человечности (то есть ширококомасштабного или систематического нападения на любых гражданских лиц⁵⁰).

Предубеждения в отчетности

Предубеждения в отчетности проявляются тогда, когда частота упоминания действий, результатов или характеристик не отражает их реального распространения или степени, в которой та или иная характеристика присуща определенной категории людей⁵¹. Предубеждения в отчетности отличаются от ошибок отбора тем, что в данном понятии на первый план выходит степень представленности в имеющихся данных, а не способ их получения. Патрик Болл и Меган Прайс отмечают: «Если термин «ошибка отбора» отсылает нас к путям определения того, какие события учитывать в процессе данных, то термин «предубеждения в отчетности» относится к ситуациям, когда одни факты замалчиваются, а другие выходят на первый план в результате действий и решений свидетелей и лиц, проводящих очный опрос»⁵². Предубеждения в отчетности могут возникать в ситуациях, когда люди документально фиксируют только те обстоятельства, которые являются необычными или особенно запоминающимися для них, подразумевая, что обычные обстоятельства и так очевидны. Кроме того, этот тип предубеждений появляется из-за того, что «более доступные данные обычно учитываются и анализируются чаще других, вследствие чего возникает предубеждение в отчетности — ведь чем сложнее получить информацию, тем меньше вероятность ее попадания в набор данных»⁵³. Например, сведения из источников на дари или пушту, имеющие значение для предварительной экспертизы событий в Афганистане, найти гораздо сложнее, чем сведения из источников на английском языке (и оттого первые реже цитируются

49 ICTY, *Prosecutor v. Duško Tadić*, Case No. IT-94-1-A, Decision on the Defence Motion for Interlocutory Appeal on Jurisdiction (Appeals Chamber), 2 October 1995, para. 70.

50 Римский статут (примечание 19 выше), ст. 7.

51 Eirini Ntoutsi *et al.*, “Bias in Data-Driven Artificial Intelligence Systems — An Introductory Survey”, *Data Mining and Knowledge Discovery*, 2019, p. 4, доступно по адресу: <https://bit.ly/3sCECmT>. См. также: Jonathan Gordon and Benjamin Van Durme, “Reporting Bias and Knowledge Acquisition”, *Proceedings of the 2013 Workshop on Automated Knowledge Base Construction*, 2013, p. 25, доступно по адресу: <https://bit.ly/2LXoD2a> (где приводится общий анализ того, как предубеждения в отчетности действуют в системах искусственного интеллекта).

52 M. Price and P. Ball (примечание 45 выше), n. 4.

53 S. K. Katyal (примечание 25 выше), p. 72.

в запросе КП о возбуждении расследования в соответствии со статьей 15 Римского статута⁵⁴).

По сравнению с другими типами предубеждения в отчетности представляют собой более серьезную проблему в сфере больших данных. К примеру, модели МО для прогнозного планирования охраны правопорядка ориентируются на информацию о том, откуда поступили сведения об убийствах, а не о том, где они были совершены. Если о преступлениях, совершаемых одной группой лиц, сообщается чаще, чем о совершаемых другими группами, модели МО, предназначенные для прогнозирования преступной активности, очевидно, будут оценивать эту группу лиц на основе предубеждения⁵⁵. В контексте КСП предубеждения в отчетности часто приводят к искажению прогнозов МО по отношению к периферийным точкам спектра⁵⁶. Так, в 2008 году в eBay отчитались, что 99% отзывов пользователей были положительными. Однако это не означает, что компании удалось добиться больших успехов в улучшении пользовательского опыта; это скорее свидетельствует о том, что пользователи eBay с гораздо меньшим желанием оставляли сообщения об отрицательном, чем о положительном опыте⁵⁷. Поэтому рассмотренный в отчете контент изначально содержал в себе предубеждения.

С учетом того, что модели МО и другие наборы больших данных, используемые для аналитики, достаточно часто основаны на КСП, эти предубеждения могут также быть обусловлены демографической структурой того множества людей, которые размещают информацию в интернете⁵⁸. Так происходит из-за того, что «большие данные в гораздо большей степени отражают сведения о наличии или присутствии чего-либо, а не об отсутствии»⁵⁹. Рикардо Баэса-Ятес объясняет: «Модели доступа к интернету и его использования коррелируют с предубеждениями, обусловленными образовательными, экономическими, технологическими и другими характеристиками, из-за чего возникает «эффект домино» в распространении предубеждений посредством веб-контента и ссылок»⁶⁰. Даже при том, что число активных пользователей Facebook (к примеру) очень велико, этой социальной сетью пользуются не все люди. Аналогичным образом, хотя название социальной сети Twitter сегодня знают почти в каждом доме, число пользователей, активно размещающих в ней сообщения, все еще относительно

54 ICC, *Situation in the Islamic Republic of Afghanistan*, Case No. ICC-02/17, Request for Authorisation of an Investigation Pursuant to Article 15 (Pre-Trial Chamber II), 20 November 2017.

55 Randy Rieland, "Artificial Intelligence Is Now Used to Predict Crime. But Is It Biased?", *Smithsonian Magazine*, 5 March 2018, доступно по адресу: <https://bit.ly/2HHg2Pf>.

56 Hongyu Chen, Zhiqiang Zheng and Yasin Ceran, "De-Biasing the Reporting Bias in Social Media Analytics", *Production and Operations Management*, Vol. 25, No. 5, 2015, p. 849.

57 Chrysanthos Dellarocas and Charles A. Wood, "The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias", *Management Science*, Vol. 54, No. 3, 2008, p. 460.

58 Berkeley Protocol on Digital Open Source Investigations, HR/PUB/20/2, 2020 (Berkeley Protocol), pp. 11, 46, 55.

59 Mick P. Couper, "Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys", *Survey Research Methods*, Vol. 7, No. 3, 2013, pp. 145, 147.

60 Ricardo Baeza-Yates, "Bias on the Web", *Communications of the ACM*, Vol. 61, No. 6, 2018, p. 54.

невелико и формируется в основном из отдельных категорий людей (22% населения США, 10% из которых публикуют 80% всех «твитов»)⁶¹. Другими словами, в социальных сетях следует отличать *производителей* контента от *потребителей* контента — структура первой группы может отличаться от структуры второй, и ни одна из них при этом не будет соответствовать по структуре населению в целом. Исследуя сообщения в Twitter, вы скорее будете изучать мнение представителей определенного слоя общества, а не мнение населения в целом.

Давайте предположим, что следователи КП создали модель МО для определения того, на какие преступления в Афганистане стоит обратить больше внимания, взяв за основу КСП, так как проведение очного расследования в этой стране затруднено или небезопасно. В такой ситуации решения следствия совершенно точно будут приниматься предвзято ввиду лингвистического неравновесия, которое влияет на онлайн-контент. В частности, согласно оценкам, английский является языком более 30% всех веб-сайтов в интернете, при том что носители английского языка как родного составляют лишь около 5% населения мира. Менее чем 0,1% интернет-контента публикуется на пушту, и лишь 3% — на дари, — а именно эти языки наиболее распространены на территории Афганистана⁶².

Еще одним дополнительным недостатком интернет-контента является ограниченность доступа к Глобальной сети для определенной доли населения⁶³. Исследования онлайн-контента изначально не позволяют учесть всю долю населения, которая не пользуется интернетом, и всех тех, кто по той или иной причине не размещает контент онлайн. В контексте КСП наиболее очевидной ситуацией возникновения предубеждений в отчетности является работа с данными мобильных телефонов. Несмотря на повсеместную распространенность мобильных устройств в определенных регионах мира, далеко не у всех людей на планете есть мобильный телефон, и уж тем более — смартфон. Так, например, в Африке к югу от Сахары достаточно высок уровень распространенности мобильных телефонов, но в то же время распространенность смартфонов здесь ниже, чем в любом другом географическом регионе⁶⁴. В ходе обследования, проведенного в 2018 году Исследовательским центром Пью, выяснилось, что «аналогично доле пользователей интернета доля людей, имеющих смартфон, в каждой обследованной стране варьируется в зависимости от возраста и уровня

61 Stefan Wojcik and Adam Hughes, “Sizing Up Twitter Users”, Pew Research Center, 24 April 2019, доступно по адресу: <https://pewrsr.ch/38TfNeD>.

62 Holly Young, “The Digital Language Divide”, *The Guardian*, доступно по адресу: <https://bit.ly/2Kn116q>; Web Technology Surveys, “Usage Statistics of Persian for Websites”, доступно по адресу: <https://bit.ly/2YN4DCk>; Web Technology Surveys, “Usage Statistics of Pushto, Pashto for Websites”, доступно по адресу: <https://bit.ly/3oNtVuz>.

63 Jill A. Dever, Ann Rafferty and Richard Valliant, “Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias?”, *Survey Research Methods*, Vol. 2, No. 2, 2008, p. 47.

64 Laura Silver and Courtney Johnson, “Majorities in Sub-Saharan Africa Own Mobile Phones, but Smartphone Adoption Is Modest”, Pew Research Center, 9 October 2018, доступно по адресу: <https://pewrsr.ch/3nVR6mj>.

образования респондентов»⁶⁵. Более того, на распространенность смартфонов значительно влияет уровень личного дохода, а в некоторых случаях — пол, так как во многих странах женщины гораздо реже владеют смартфонами, чем мужчины⁶⁶.

Вернемся к гипотетическому примеру расследования событий в Афганистане и вспомним, что один политический аналитик выяснил, что в 2008 году из общего числа нападений повстанцев, повлекших за собой хотя бы одну человеческую жертву, только 30% нашли отражение в международных новостях, и «таким образом, из информации СМИ об Афганистане можно получить представление менее чем о трети от общего реального числа случаев проявления насилия»⁶⁷. Одной из причин такого расхождения информации с действительностью было то, что сообщения о насилии поступали «в подавляющем большинстве из мест, охваченных зоной покрытия сотовой связи» — таким образом, именно от охвата сотовой связью зависело, будет ли то или иное происшествие отражено в СМИ⁶⁸. Научные исследования также говорят о том, что присутствие военного контингента США значительно влияет на интенсивность освещения событий в стране в новостях и распространения информации о них в СМИ⁶⁹. Например, благодаря нахождению в Афганистане военнослужащих из большого количества западных стран, входящих в международную коалицию, менее масштабные события с меньшим количеством жертв здесь освещаются активнее, чем эпизоды многолетней гражданской войны в Центральноафриканской Республике, о которой очень редко пишут крупные международные СМИ⁷⁰.

С точки зрения расследования массовых зверств или нарушений МГП существует несколько сценариев, в которых могут проявляться предубеждения в отчетности. Преступления, влекущие за собой большое число жертв или совершаемые особо бесчеловечным образом, оставляют исключительно глубокий информационный след; предположительно менее серьезные (то есть с малым числом жертв) преступления же, напротив, редко получают широкое освещение, особенно если происходят в рамках жестокого и продолжительного вооруженного конфликта. Это говорит о том, что информация о таких событиях из открытых источников потенциально может вводить в заблуждение. Как отмечено в докладе Федеральной торговой комиссии США о больших данных, «большие данные могут быть

65 Jacob Poushter, Caldwell Bishop and Hanyu Chwe, “Smartphone Ownership on the Rise in Emerging Economies”, Pew Research Center, 19 June 2018, доступно по адресу: <https://pewrsr.ch/2Negkjr>.

66 Pew Research Center, “Mobile Fact Sheet”, 12 June 2019, доступно по адресу: <https://pewrsr.ch/2LMq9El>.

67 Nils B. Weidmann, “A Closer Look at Reporting Bias in Conflict Event Data”, *American Journal of Political Science*, Vol. 60, No. 1, 2015, p. 211.

68 Ibid., p. 217.

69 Timothy M. Jones, Peter Van Aelst and Rens Vliegthart, “Foreign Nation Visibility in U.S. News Coverage: A Longitudinal Analysis (1950–2006)”, *Communication Research*, Vol. 40, No. 3, 2013, p. 417.

70 N. B. Weidmann (примечание 67 выше), p. 216 and Appendix D.

весьма эффективны при демонстрации взаимосвязей, однако следует прежде всего понимать, что взаимосвязь и причинно-следственная связь — разные вещи. И действительно, при анализе достаточно крупных наборов данных всегда можно найти определенное количество бессмысленных взаимосвязей»⁷¹.

По причине включенных в набор данных предубеждений в отчетности модель МО, разработанная для выявления закономерностей в преступлениях или нападениях повстанцев, генерирует вводящие в заблуждение выводы о распространенности преступлений или нападений или об их общих характеристиках. Болл и Прайс пронизательно замечают, что

об убийствах в городах сообщается практически всегда, а убийства в сельской местности документируются достаточно редко. Следовательно, вероятность отражения события в отчетности зависит от того, где оно произошло. В результате анализ, выполненный на таких данных без учета предубеждений, всегда будет указывать на то, что насилие в основном происходит в городах⁷².

В связи с этим, подобные предубеждения могут быть причиной формирования ложных представлений о потенциальных закономерностях и моделях преступлений или нападений. Между тем такие выводы имеют большое значение при оценке того, совершаются ли преступления в рамках плана или политики⁷³, были ли деяния обвиняемого умышленными и случайными⁷⁴, или того, имеет ли отношение то или иное деяние к определенному вооруженному конфликту.

Наибольший ущерб могут причинять предубеждения в отчетности, связанной с традиционно не учитываемыми в полной мере преступлениями или нарушениями МГП, такими как преступления на сексуальной и гендерной почве (ПСГП)⁷⁵. Как отмечается в одном из докладов Генерального секретаря ООН, сексуальное насилие, связанное с конфликтом, обычно не полностью фиксируется в отчетности вследствие «запугивания и стигматизации, а также ограничений на доступ для персонала Организации Объединенных Наций»⁷⁶. Неполный учет сексуального насилия повлиял

71 US Federal Trade Commission, *Big Data: A Tool for Inclusion or Exclusion?*, 2016, p. 9, доступно по адресу: <https://bit.ly/31Or102>. См. также: Martin Frické, “Big Data and Its Epistemology”, *Journal of the Association for Information Science and Technology*, Vol. 66, No. 4, 2015, p. 659.

72 M. Price and P. Ball (примечание 45 выше), pp. 10–11.

73 См., например: ICTY, *Prosecutor v. Nikola Šainović et al.*, Case No. IT-05-87-A, Judgment (Appeals Chamber), 23 January 2014, paras 614–634 (где был оставлен в силе вывод Судебной камеры о том, что наличие общего плана по перемещению албанского населения Косово подтверждает «явно прослеживаемой схемой» насильственного выселения).

74 См., например: *ibid.*, paras 988, 1784; *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Judgment Pursuant to Article 74 of the Statute (Trial Chamber VII), 19 October 2016, paras 702, 707 (где в целях оценки намерений обвиняемых совершить преступление отмечается систематический характер их действий).

75 Berkeley Protocol (примечание 58 выше), p. 57.

76 *Генеральный секретарь ООН. Сексуальное насилие, связанное с конфликтом: Доклад Генерального секретаря*, док. ООН S/2019/280, 29 марта 2019 г. (Доклад ГС ООН о сексуальном насилии),

как минимум на три расследуемые МУС в настоящее время ситуации в разных странах, в частности в Афганистане⁷⁷, Центральноафриканской Республике⁷⁸ и Ливии⁷⁹. В подготовленном Прокурором МУС Программном документе о преступлениях на сексуальной и гендерной почве признается существование «специфических проблем», с которыми сопряжено расследование ПСГП, в том числе «недостаточного или нулевого учета актов сексуального насилия в силу социальных, культурных или религиозных факторов» и «обусловленной этим нехватки оперативно доступных улик»⁸⁰. Устоявшаяся практика неполного учета ПСГП создает условия, в которых наборы данных о международных преступлениях, особенно представленные в онлайн-форматах, будут давать заведомо искаженное представление о распространенности ПСГП во время того или иного конфликта. Это скажется и на всех моделях МО и попытках анализа больших данных, в которых будут использоваться такие наборы. В результате МО и анализ больших данных могут усугубить описанные выше предубеждения.

Ошибка групповой атрибуции

Ошибка групповой атрибуции — это тенденция к экстраполяции фактов, правдивых в отношении нескольких лиц, на всю группу, к которой они принадлежат. Представим, например, что создается модель МО для выявления наиболее подходящих кандидатов на позиции в КП. В процессе разработки авторы по умолчанию устанавливают, что «лучшие» кандидаты — это обладатели докторской степени, полученной в университетах стран Западной Европы, и опыта стажировки в МУС; такой вывод авторы делают исходя лишь из того, что эти характеристики присущи ряду успешных сотрудников. Созданная модель будет содержать ошибку групповой атрибуции из-за того, что столь же или более квалифицированные специалисты, не обладающие указанными характеристиками, с ее точки зрения будут считаться менее значимыми кандидатами.

п. 11. См. также: World Health Organization, *Global and Regional Estimates of Violence against Women: Prevalence and Health Effects of Intimate Partner Violence and Non-Partner Sexual Violence*, 20 October 2013, доступно по адресу: <https://bit.ly/3oXrFlp>; Iness Ba and Rajinder S. Bophal, “Physical, Mental and Social Consequences in Civilians Who Have Experienced War-Related Sexual Violence: A Systematic Review (1981–2014)”, *Public Health*, Vol. 142, 10 September 2016; Gerald Schneider, Lilli Banholzer and Laura Albarracin, “Ordered Rape: A Principal–Agent Analysis of Wartime Sexual Violence in the DR Congo”, *Violence Against Women*, Vol. 21, No. 11, 2015; Tia Palermo, Jennifer Bleck and Amber Peterman, “Tip of the Iceberg: Reporting and Gender-Based Violence in Developing Countries”, *American Journal of Epidemiology*, Vol. 179, No. 5, 2014.

77 Доклад ГСООН о сексуальном насилии (примечание 76 выше), пп. 31–34.

78 Там же, пп. 35–39. См. также: *Группа экспертов ООН по Центральноафриканской Республике*. Заключительный доклад Группы экспертов по Центральноафриканской Республике, мандат которой был продлен во исполнение резолюции 2399 (2018) Совета Безопасности, док. ООН S/2018/1119, 14 декабря 2018 г., пп. 164–167; Phuong N. Pham, Mychelle Balthazard and Patrick Vinck, “Assessment of Efforts to Hold Perpetrators of Conflict-related Sexual Violence Accountable in Central African Republic”, *Journal of International Criminal Justice*, Vol. 18, No. 2, 2020, pp. 394–395.

79 Доклад ГСООН о сексуальном насилии (примечание 76 выше), пп. 54–59.

80 ICC, OTP, *Policy Paper on Sexual and Gender-Based Crimes*, June 2014, доступно по адресу: <https://bit.ly/3in5nHk> (OTP Policy Paper on SGBC), para. 50.

Ошибка групповой атрибуции в основном проявляется в виде двух эффектов: эффекта «своей» группы и эффекта «чужой» группы. Эффект «своей» группы заключается в тенденции к более положительному восприятию членов группы, к которой принадлежите вы сами, или людей, у которых есть те же черты, что и у вас⁸¹. Эффект «чужой» группы, напротив, отражается в тенденции стереотипно воспринимать отдельных членов группы, к которой вы сами не принадлежите, или считать их характеристики более универсальными для этой группы⁸². В связи с этим, мы более тонко различаем черты членов «своей» группы, нежели черты членов иных групп⁸³.

С точки зрения МГП ошибку групповой атрибуции следует учитывать в ситуациях, когда МО используется для определения или прогнозирования того, является ли тот или иной человек комбатантом, в рамках процесса прицеливания. В 2009 году ученые из Нориджского университета — высшего военного училища — провели исследование, в рамках которых курсантам-мужчинам было поручено быстро принимать решение об открытии огня при непродолжительном показе огнестрельного оружия на экране компьютера⁸⁴. Курсанты быстрее и точнее реагировали на появление огнестрельного оружия на экране после демонстрации фотографий мужчин в традиционной для Ближнего Востока одежде, а также делали больше ложноположительных ошибок, когда изображения оружия чередовались с подобными фотографиями. Так происходило потому, что у курсантов, выросших в период после террористических актов 11 сентября 2001 года, из-за которых США сосредоточили деятельность своих вооруженных сил в Ираке и Афганистане, сформировались стереотипы о мужчинах ближневосточной наружности, особенно носящих традиционные халаты и тюрбаны. В данном случае стереотипы диктуют, что такие мужчины с большей вероятностью являются террористами или неприятельскими комбатантами⁸⁵. Реакция моделей МО аналогична реакции курсантов. Модели, разработанные на основе наборов данных, которые ориентированы только на одну группу, значительно чаще будут обуславливать ошибки при прицеливании в силу стереотипного отрицательного восприятия такой группы.

81 См.: S. K. Katyal (примечание 25 выше), pp. 80–81, где цитируется Michael J. Bernstein, Steven G. Young and Kurt Hugenberg, “The Cross-Category Effect: Mere Social Categorization Is Sufficient to Elicit an Own-Group Bias in Face Recognition”, *Psychological Science*, Vol. 18, No. 8, 2007.

82 См.: S. K. Katyal (примечание 25 выше), p. 81, где цитируется S. Alex Haslam, Penny J. Oakes and John C. Turner, “Social Identity, Self-Categorization, and the Perceived Homogeneity of Ingroups and Outgroups: The Interaction Between Social Motivation and Cognition”, в Richard M. Sorrentino and Edward T. Higgins (eds), *Handbook of Motivation and Cognition: The Interpersonal Context*, Vol. 3, Guilford Press, New York, 1996.

83 Donald M. Taylor and Janet R. Doria, “Self-Serving and Group-Serving Bias in Attribution”, *Journal of Social Psychology*, Vol. 113, No. 2, 1981.

84 Kevin K. Fleming, Carole L. Bandy and Matthew O. Kimble, “Decisions to Shoot in a Weapon Identification Task: The Influence of Cultural Stereotypes and Perceived Threat on False Positive Errors”, *Social Neuroscience*, Vol. 5, No. 2, 2010.

85 Ibid., pp. 206, 219. См. также: В. Keith Payne and Joshua Correll, “Race, Weapons, and the Perception of Threat”, в Bertram Gawronski (ed.), *Advances in Experimental Social Psychology*, Vol. 62, Elsevier, Amsterdam, 2020, Chap. 1.

Предубеждения этого типа также обнаруживаются в решениях судебных органов тех стран, где наблюдается значительная поляризация общества. Так, исследование решений, вынесенных в Израиле с 2000 по 2004 год судами малых исков (мировыми судами), где дела случайным образом распределялись между судьями арабского и еврейского происхождения, выявило, что судьи в среднем на 17–20% чаще принимали иски к производству, если истец принадлежал к той же этнической группе, что и они⁸⁶. Авторы этого исследования также пришли к выводу о том, что распространенность эффекта «своей» группы была выше в районах, недавно затронутых террористическими актами, а также о том, что «в районах, где в недавнем прошлом не наблюдалось относительно высокой межэтнической напряженности, распространенность таких предубеждений была существенно ниже»⁸⁷. Аналогичное исследование, в котором изучалась работа более 100 судей в США, позволило обнаружить, что судьи также находились под влиянием эффекта «своей» группы⁸⁸. Судьи, в большей степени симпатизирующие белокожим ответчикам/обвиняемым, выносили более суровые решения в отношении чернокожих ответчиков/обвиняемых⁸⁹. Аналогичным образом симпатизирующие чернокожим ответчикам/обвиняемым судьи были более снисходительны к ним, нежели к белокожим⁹⁰. Ряд других недавних научных работ продемонстрировал, что судьи систематически предвзято принимают отрицательные решения в отношении членов «своей» группы — при рассмотрении дел против белокожего обвиняемого белокожим судьей и против чернокожего обвиняемого — чернокожим судьей назначаемый срок лишения свободы был на 14% дольше, чем при вынесении приговора обвиняемому из «чужой» группы⁹¹.

Несложно представить себе обстоятельства, в которых ошибки групповой атрибуции могут повлиять на исход дел, слушаемых в МУС. Как отмечалось в последнее время, например в Независимом экспертном обзоре МУС и Римского статута, «многие из собеседников экспертов, в том числе и сами судьи, говорили о чрезвычайной «привязанности» отдельных судей к правовым системам их родных стран, будь то в вопросах общего или гражданского права, что было одной из причин отсутствия согласованности практических методов разных Палат между собой»⁹². В некоторых

86 Moses Shayo and Asaf Zussman, “Judicial Ingroup Bias in the Shadow of Terrorism”, *Quarterly Journal of Economics*, Vol. 126, No. 3, 2011, p. 1447.

87 *Ibid.*, p. 1483.

88 Jeffrey J. Rachlinski, Sheri Lynn Johnson, Andrew J. Wistrich and Chris Guthrie, “Does Unconscious Bias Affect Trial Judges?”, *Notre Dame Law Review*, Vol. 84, No. 3, 2009, pp. 1225–1226.

89 *Ibid.*, p. 1223.

90 *Ibid.* При этом см. p. 1223 (где демонстрируется, что при случаях явного манипулировании расовой принадлежностью судьи способны сохранять объективное отношение к обвиняемым).

91 Jeff Guo, “Researchers Have Discovered a New and Surprising Racial Bias in the Criminal Justice System”, *Washington Post*, 24 February 2016, доступно по адресу: <https://wapo.st/37Nz0hR>; Briggs Depew, Ozkan Eren and Naci Mocan, “Judges, Juveniles and In-Group Bias”, *Journal of Law and Economics*, Vol. 60, No. 2, 2017.

92 *Independent Expert Review of the International Criminal Court and the Rome Statute System: Final Report*, 30 September 2020 (IER Report), para. 632, доступно по адресу: <https://bit.ly/2XSkA9Z>.

случаях судьи МУС столь твердо придерживались таких соображений, что выносили решения на основании правовой традиции своих стран, даже при явном их несоответствии правовым актам Суда. Например, в одной из своих университетских лекций судья Марк Перрен де Бришамбо, гражданин Франции, отметил, что вместе с другими членами своей коллегии, специализирующимися на гражданском праве, он решил не удовлетворять ни одну из промежуточных апелляций, поданных при рассмотрении дела Бембы и др., несмотря на то, что такие апелляции предусмотрены Римским статутом. Бришамбо аргументировал такое решение тем, что промежуточные апелляции обыкновенно не разрешены в странах романо-германской правовой системы: «В деле Бембы и др. мы выступали в качестве специалистов по соответствующей правовой системе. Мы заявили, что механизм промежуточных апелляций неприемлем, и потому возможность их подачи не будет принята во внимание»⁹³.

Мнение Бришамбо можно рассматривать как проявление ошибки групповой атрибуции в пользу «своей» группы, так как он предпочел сделать выбор в пользу правил судопроизводства, основанных на правовой традиции, к которой он сам принадлежит, пренебрегая тем фактом, что такое решение противоречит другим правовым традициям или даже напрямую формулировкам Римского статута. При разработке модели МО, предназначенной для прогнозирования судебных решений, такой как существующая модель прогнозирования решений Европейского суда по правам человека⁹⁴, ошибки групповой атрибуции следует обязательно учитывать.

Предубеждения в пользу автоматизированных систем

Предубеждения в пользу автоматизированных систем — это склонность к предпочтению выводов, сделанных автоматизированными или компьютерными системами, тем выводам, которые были получены с использованием неавтоматизированных систем, без учета частоты совершения ошибок в обоих типах систем⁹⁵. Было отмечено, что «предубеждение в пользу автоматизированных систем по сути придает предложенному компьютерной программой варианту статус авторитетного окончательного решения»⁹⁶.

93 Marc Perrin de Brichambaut, “ICC Statute Article 68”, Peking University Law School, Beijing, 17 May 2017, p. 9, доступно по адресу: <https://bit.ly/35SIYg4>.

94 Masha Medvedeva, Michel Vols and Martijn Wieling, “Using Machine Learning to Predict Decisions of the European Court of Human Rights”, *Artificial Intelligence and Law*, Vol. 28, 2020; Conor O’Sullivan and Joeran Beel, “Predicting the Outcome of Judicial Decisions Made by the European Court of Human Rights”, *27th AIAI Irish Conference on Artificial Intelligence and Cognitive Science*, 2019, доступно по адресу: <https://bit.ly/3nXEBGO>.

95 Linda J. Skitka, Kathleen Mosier, Mark Burdick and Bonnie Rosenblatt, “Automation Bias and Errors: Are Crews Better Than Individuals?”, *International Journal of Aviation Psychology*, Vol. 10, No. 1, 2000, p. 86; Ric Simmons, “Big Data, Machine Judges, and the Legitimacy of the Criminal Justice System”, *UC Davis Law Review*, Vol. 52, No. 2, 2018, pp. 1109–1110; Mary L. Cummings, “Automation and Accountability in Decision Support System Interface Design”, *Journal of Technology Studies*, Vol. 32, No. 1, 2006, p. 25.

96 Danielle K. Citron, “Technological Due Process”, *Washington University Law Review*, Vol. 85, No. 6, 2008, p. 1272.

В контексте МО данный тип предубеждений в основном проявляется, когда существуют различия между реально преследуемой целью и представлением автоматизированной системы об этих целях и любых релевантных ограничениях. Если в алгоритме не учтены (среди прочих) социальные, культурные или политические факторы или пользователь не способен выявить свойственные алгоритму ограничения, это ведет к появлению предубеждений в пользу автоматизированных систем.

Предубеждения в пользу автоматизированных систем представляют значительный риск при проведении оценки с точки зрения МГП. К примеру, в контексте автономных систем оружия и систем «поддержки решений», применяемых в процессе прицеливания, эти предубеждения становятся причиной чрезмерной уверенности людей в функциональных способностях таких систем, степень которой доходит до вверения «задачи по моральному обоснованию и обеспечению подотчетности машинам как предположительно легитимному авторитетному субъекту»⁹⁷. Более того, предубеждения в пользу автоматизированных систем могут приводить к аналогичной проблеме и в сфере МГП, если считается, что модель МО с большей вероятностью будет правильно выделять комбатантов и захватывать их как цели, нежели оператор-человек. Чем сложнее модель МО, тем чаще проявляется склонность доверять машинам и не вмешиваться даже тогда, когда становится понятным, что машина совершила ошибку. Так происходит из-за веры в надежность более сложных машин⁹⁸.

В правоприменительной сфере предубеждения в пользу автоматизированных систем приводят к тому, что юристы и следователи охотнее соглашались с заключениями, полученными от моделей МО или других инструментов анализа больших данных, а не от систем с человеком в контуре управления, пренебрегая риском наличия предубеждений в таких заключениях. Влияние предубеждений в пользу автоматизированных систем имеет дополнительный эффект: возникает предположение о том, что методы поиска и извлечения данных в интернете в целом или в социальных сетях в частности позволяют обнаружить достоверные доказательства быстрее и действеннее, чем традиционные методы расследования. До настоящего времени эмпирические сведения, которые бы подтверждали это предположение, отсутствуют. К примеру, в Соединенных Штатах НКО ProPublica изучила разработанную институтом Northpointe систему составления профилей заключенных в контексте замены тюремного заключения на альтернативные санкции (COMPAS) — это система оценки рисков в рамках уголовных дел, которую применяют при вынесении приговоров

97 ICRC, *Ethics and Autonomous Weapon Systems: An Ethical Basis for Human Control?*, Geneva, 3 April 2018, p. 14, доступно по адресу: <https://bit.ly/3ioj3C5>.

98 Chantal Grut, “The Challenge of Autonomous Lethal Robotics to International Humanitarian Law”, *Journal of Conflict and Security Law*, Vol. 18, No. 1, 2013, p. 19. См. также: Shin-Shin Hua, “Machine Learning Weapons and International Humanitarian Law: Rethinking Meaningful Human Control”, *Georgetown Journal of International Law*, Vol. 51, No. 1, 2019, p. 141.

и решений по поводу условно-досрочного освобождения⁹⁹. Исследование ProPublica выявило, что алгоритм COMPAS предвзято относился к обвиняемым разных рас, отмечая вероятность рецидива у чернокожих обвиняемых как «высокую», а у белокожих — как «низкую». Однако это не помешало Верховному суду штата Висконсин одобрить использование алгоритма COMPAS в деле «Государство против Лумиса», не предложив какой-либо реальной надлежащей процедуры¹⁰⁰, что, по мнению многих экспертов, вызвано предубеждением суда в пользу автоматизированных систем¹⁰¹.

Предубеждения в пользу автоматизированных систем пока не являются значительной проблемой в работе МУС или других учреждений в сфере МУП, так как Суд, насколько известно автору настоящей статьи, не прибегает к использованию автоматизированных систем или МО при принятии решений. Однако по мере того, как международные учреждения будут рассматривать возможность внедрения новых технологий, в том числе МО, в свою деятельность, можно ожидать возникновения аналогичных случаев, если не будут реализованы надлежащие меры предосторожности и механизмы контроля.

Правовые последствия предубеждений с точки зрения МГП

Как было продемонстрировано выше, модели МО могут использоваться во время вооруженных конфликтов различными способами. Технологии МО могут быть интегрированы в сложное автономное вооружение, в том числе в системы перехвата неуправляемых ракет, артиллерийских снарядов и минометных боеприпасов (C-RAM), которые могут обладать определенной степенью автономности при обнаружении, отслеживании, выборе и атаке целей¹⁰². Более того, модели МО применяются для поддержки принятия человеком решений о месте и времени нападения. Объединенный центр искусственного интеллекта США, к примеру, планирует использовать модели МО для повышения осведомленности об обстановке и эффективности принятия решений, повышения уровня безопасности оборудования и реализации прогнозного обслуживания и снабжения¹⁰³. Такое применение МО сопряжено с возможными последствиями с точки зрения МГП; в следующих двух разделах освещены два вида таких последствий.

99 Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, “Machine Bias: There’s Software Used across the Country to Predict Future Criminals. And It’s Biased against Blacks”, *ProPublica*, 23 May 2016, доступно по адресу: <https://bit.ly/39GHiHK>.

100 Wisconsin Supreme Court, *State v. Loomis*, 881 N.W.2d 749, 13 July 2016, pp. 770–771.

101 Aleš Završnik, “Criminal Justice, Artificial Intelligence Systems, and Human Rights”, *ERA Forum*, Vol. 20, No. 4, 2020; Katherine Freeman, “Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in *State v. Loomis*”, *North Carolina Journal of Law & Technology*, Vol. 18, No. 5, 2016, pp. 97–98.

102 ICRC, *Autonomous Weapon Systems: Implications of Increasing Autonomy in the Critical Functions of Weapons. Expert Meeting*, Geneva, March 2016, p. 10, доступно по адресу: <https://bit.ly/35VHscW>.

103 DoD, *Summary of the 2018 Department of Defense Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*, 2018, p. 7, доступно по адресу: <https://bit.ly/2LKOSZN>.

Проверка вооружений

Право комбатантов выбирать средства и методы войны ограничено рядом норм МГП, в том числе международными договорами, запрещающими использовать конкретные виды оружия¹⁰⁴. Такие нормы дополняются принципами МГП относительно новых видов оружия, средств или методов ведения войны; эти принципы предназначены для того, чтобы предотвратить такое применение новых технологий в условиях боевых действий, которое бы нарушало международное право, и чтобы гарантировать установление соответствия таких технологий международному гуманитарному праву до их разработки, приобретения или включения в арсенал государства другими способами¹⁰⁵. В частности, согласно статье 36 Дополнительного протокола I «при изучении, разработке, приобретении или принятии на вооружение новых видов оружия, средств или методов ведения войны» все государства-участники должны определить, подпадает ли использование этого оружия под запрещения, содержащиеся в ДП I «или в каких-либо других нормах международного права», применяемых к соответствующему государству.

По всей видимости, статья 36 применима к системам оружия, использующим модели МО для улучшения способности систем прицеливания проводить различие между гражданскими лицами, дружественными солдатами и неприятельскими комбатантами или даже для обоснования принимаемых решений по поводу вооруженных нападений, так как подобные системы практически во всех случаях можно охарактеризовать как «методы» или «средства» ведения войны. Как отмечено в Комментарий к ДП I, оба термина используются для того, чтобы «обозначать виды оружия в максимально широком смысле, а также способы их применения»¹⁰⁶. Термин «средства ведения войны» «в общем обозначает действия по применению оружия»¹⁰⁷, в отличие от термина «методы ведения войны», который обозначает «то, как используется оружие»¹⁰⁸. В этом отношении можно сказать, что статья 36 по существу охватывает не только «новые» виды оружия в техническом смысле, но и, как указано в Руководстве МККК, «способы, которыми эти виды оружия должны использоваться в соответствии с воен-

104 В ДП I содержатся положения, которые устанавливают ограничения на использование различных видов оружия, средств и методов ведения войны и защищают гражданское население от последствий боевых действий. См. Дополнительный протокол I к Женевским конвенциям от 12 августа 1949 г., касающийся защиты жертв международных вооруженных конфликтов, 8 июня 1977 г. (вступивший в силу 7 декабря 1978 г.), в частности часть III, раздел I, и часть IV, раздел I, главы I–IV.

105 МККК. Руководство по проверке соответствия нормам права новых видов оружия, средств и методов ведения войны: меры по имплементации статьи 36 Дополнительного протокола I 1977 г., Москва, 2006 (Руководство МККК по проверке новых видов оружия). С. 3–4.

106 Yves Sandoz, Christophe Swinarski and Bruno Zimmermann (eds), *Commentary on the Additional Protocols*, ICRC, Geneva, 1987 (ICRC Commentary on APs), para. 1402.

107 Ibid., para. 1957. См. также: Michael N. Schmitt, *Tallinn Manual on the International Law Applicable to Cyber Warfare*, Cambridge University Press, Cambridge, 2013, Rule 41(b).

108 ICRC Commentary on APs (примечание 106 выше), para. 1957.

ной доктриной, тактикой, правилами применения силы, инструкциями и контрмерами»¹⁰⁹. И действительно, в Комментариях к статье 36 отдельно делается акцент на том, что разработка новых видов или систем оружия ведет «к автоматизации театра боевых действий, где роль солдата становится все менее значительной», и высказывается обеспокоенность тем, «что, если человек не подчинит себе технологию, а подчинится ей, она уничтожит его»¹¹⁰.

При этом гораздо более неясным остается вопрос о том, достаточно ли широка сфера действия статьи 36, чтобы охватить системы поддержки принятия решений в боевых условиях, использующие модели МО, но при этом не задействованные напрямую в процессе прицеливания. Этот вопрос актуален в обстоятельствах, когда МО, как в случае с Объединенным центром искусственного интеллекта США, используется для повышения уровня осведомленности об обстановке, проведения прогнозного техобслуживания и снабжения или повышения уровня безопасности оборудования. Корень вопроса лежит в том, можно ли отнести такие системы к «методам ведения войны». Ответ в таких обстоятельствах, скорее всего, будет отрицательным, так как даже наиболее широкое определение «метода ведения войны» подразумевает наличие непосредственной связи с использованием оружия или его функционированием¹¹¹. Оценивать то, влияет ли модель МО на «способы, которыми эти виды оружия должны использоваться», следует на индивидуальной основе. В обстоятельствах, когда применение модели ограничивается техническим обслуживанием и снабжением, наличие такого влияния маловероятно. Напротив, модели МО, применяемые при выработке военной тактики, правил ведения боевых действий или порядка осуществления операций, очевидно, непосредственно подпадают под действие статьи 36.

Весьма вероятно, что в силу статьи 36 государства должны учитывать и распространенные человеческие предубеждения «при изучении, разработке, приобретении или принятии на вооружение» моделей МО в той степени, в которой они влияют на обычное и ожидаемое использование оружия. Обусловлено это тем, что предубеждения в моделях МО, встроенных в системы оружия или служащих для обоснования принятия решений в боевых условиях, могут привести к нарушению специальных или общих запретов в отношении видов оружия, средств и методов ведения войны, которые установлены договорным и обычным международным правом. Например, модели МО, которые применяются в системах прицеливания и при этом не учитывают возможное наличие в них человеческих предубеждений, могут приводить к нарушению установленных общим международным правом запрета на использование средств и методов ведения войны, способных нанести чрезмерные повреждения и причинить излиш-

109 Руководство МККК по проверке новых видов оружия (примечание 105 выше). С. 8–9.

110 ICRC Commentary on APs (примечание 106 выше), para. 1476.

111 Руководство МККК по проверке новых видов оружия (примечание 105 выше). С. 8–9.

ние страдания¹¹², и запрета на использование оружия, которое обладает неизбирательным действием¹¹³. И действительно, некоторые НПО выражают обеспокоенность тем, что «после окончания разработки полностью автономные системы вооружений с большой вероятностью могут попасть в распоряжение безответственно действующих государств или негосударственных вооруженных формирований, что даст последним возможность запрограммировать эти машины на неизбирательное уничтожение собственных граждан или населения территорий, контролируемых неприятелем»¹¹⁴.

Наконец, можно с уверенностью сказать, что статья 36 обязывает государства учитывать потенциальное воздействие предубеждений на нормы, касающиеся проведения различия и соразмерности. При том, что эти нормы в основном определяются в полевых условиях и в зависимости от конкретной ситуации, они релевантны и при оценке новых видов оружия «в такой степени, что его характеристики, ожидаемое применение и предсказуемые последствия применения дают возможность проверяющему органу или должностному лицу определить, можно ли будет применять это оружие на законных основаниях в определенных предсказуемых ситуациях и при определенных условиях»¹¹⁵. Принцип проведения различия требует от всех участников отличать гражданских лиц от комбатантов и гражданские объекты от военных целей¹¹⁶. Принцип соразмерности запрещает нападение на военные цели, которое «как можно ожидать, попутно повлечет за собой потери жизни среди гражданского населения, ранения гражданских и ущерб гражданским объектам или то и другое вместе, которые были бы чрезмерны по отношению к конкретному и непосредственному военному преимуществу, которое предполагается таким образом получить»¹¹⁷. Им также предусмотрена обязанность всех сторон в конфликте принимать все разумные меры предосторожности для защиты гражданских лиц и объектов¹¹⁸. В число таких мер предосторожности входит оценка риска для гражданских лиц и выбор подходящих времени и средства нападения¹¹⁹.

112 Хенкертс, Жан-Мари, и Досвальд-Бек, Луиза. Обычное международное гуманитарное право. Том 1: Нормы. МККК, 2006 (Обычное МГП), норма 70, с. 303, доступно по адресу: <https://ihl-databases.icrc.org/customary-ihl/rus/docs/v1>.

113 Там же, норма 71, с. 313; см. также норму 11, с. 48. См. также: Международный суд ООН, Консультативное заключение относительно законности угрозы ядерным оружием или его применения, 8 июля 1996 г. Док. ООН A/51/218, пп. 78, 95.

114 Bonnie Docherty, “Mind the Gap: The Lack of Accountability for Killer Robots”, *Human Rights Watch*, 9 April 2015, p. 7, доступно по адресу: <https://tinyurl.com/16fvbit4>.

115 Руководство МККК по проверке новых видов оружия (примечание 105 выше). С. 14.

116 ДП I, ст. 51(4)(b); ICTY, *Prosecutor v. Dragomir Milošević*, Case No. IT-98-29/1-A, Judgment (Appeals Chamber), 12 November 2009, para. 53; ICTY, *Prosecutor v. Stanislav Galić*, Case No. IT-98-29-A, Judgment (Appeals Chamber), 30 November 2006, para. 190; ICTY, *Prosecutor v. Tihomir Blaškić*, Case No. IT-95-14-A, Judgment (Appeals Chamber), 29 July 2004, para. 109.

117 ДП I, ст. 51(5)(b), 57(2)(iii); ICTY, *Galić* (примечание 116 выше), para. 190.

118 ДП I, ст. 57(4); см. также ст. 57(2).

119 Jean-François Quéguiner, “Precaution Under the Law Governing the Conduct of Hostilities”, *International Review of the Red Cross*, Vol. 88, No. 864, 2006, pp. 793, 797–808.

Как указано выше, если модели МО подвержены ошибке групповой атрибуции, существует серьезный риск того, что системы прицеливания, полагающиеся на такие модели, примут гражданских лиц за комбатантов в силу стереотипов, содержащихся в наборах данных, которые использовались для обучения. Ярким примером может служить активно разрабатываемая в США и большинстве стран Европейского союза методология оценки сопутствующего ущерба (МОСУ), которая служит для определения риска нанесения случайных ранений гражданским лицам и разрушения объектов гражданской инфраструктуры при нападении на цели, не подпадающие под защиту права¹²⁰. Параметры МОСУ в основном засекречены, но, по всей видимости, в ней применяется компьютерное моделирование, и вполне разумно ожидать, что использование моделей МО может повысить эффективность ее будущих версий (если они еще не используются в текущей версии)¹²¹. Если в таких моделях не будут учтены предубеждения, заставляющие систему относить определенных гражданских лиц к комбатантам или объекты гражданской инфраструктуры — к военным целям и обусловленные субъективностью наборов данных, МОСУ может недооценить размер потенциального ущерба и привести к увеличению ущерба для гражданских лиц и объектов.

С учетом известной возможности человеческих предубеждений воздействовать на модели МО эти нормы МГП требуют, чтобы разработчики таких моделей принимали во внимание распространенные человеческие предубеждения при изучении, разработке, приобретении или внедрении моделей МО. Это требование в том числе предполагает, что пользователи моделей МО, используемых в системах вооружений и процессах принятия решений относительно вооруженных нападений, должны гарантировать способность таких моделей делать выбор в пользу видов оружия, средств или методов ведения войны, позволяющих минимизировать число жертв среди гражданского населения. Так, например, в соответствии с нормами МГП относительно прицеливания, оружие, средство или метод ведения войны должны предусматривать возможность соблюсти принципы проведения различия, соразмерности и принятия мер предосторожности в контексте отдельной операции и с учетом конкретных предубеждений, содержащихся в данных, которые использовались для обучения программы МО. На практике это означает, что человеку можно поручить испытание и мониторинг функционирования системы оружия или программирования, чтобы снизить воздействие таких предубеждений на любое принимаемое решение. В настоящее время МККК придерживается примечательной позиции, согласно которой модели МО практически неизбежно обуславливают «характерную непредсказуемость, необъяснимость и наличие предубежде-

120 Michael N. Schmitt, "Targeting and International Humanitarian Law in Afghanistan", *International Law Studies*, Vol. 85, No. 1, 2009, p. 311. См. также: ICRC, *Autonomous Weapon Systems: Technical, Military, Legal and Humanitarian Aspects*, Geneva, March 2014, p. 83, доступно по адресу: <https://bit.ly/3c7h1F1>; Maura Riley, "Killer Instinct: Lethal Autonomous Weapons in the Modern Battle Landscape", *Texas Law Review*, Vol. 95, 2017, pp. 33–34, доступно по адресу: <https://bit.ly/3iFAsGp>.

121 M. N. Schmitt (примечание 120 выше).

ний» в структуре и при использовании любых систем, куда они интегрированы¹²². В свете такого положения дел появляется вероятность того, что испытания и мониторинг в принципе (независимо от их интенсивности) не могут гарантировать, что система, основанная на МО, успешно пройдет проверку вооружений. Наконец, правовые проверки не устраняют необходимости дискуссий по поводу применения МГП к разнообразным способам использования МО в военном деле. Исходя из того, что правовые проверки сами по себе не позволяют в достаточной степени учесть потенциальное воздействие предубеждений на работу моделей МО, применяемых в военных целях (особенно при существующих технических ограничениях), всегда будет возникать вопрос, стоит ли сразу отказаться от использования таких систем и стоит ли ввести дополнительные механизмы контроля, такие как мониторинг со стороны человека, чтобы предотвратить непредвиденный ущерб.

Наконец, даже с учетом того, что в статье 36 ДП I не описан конкретный механизм таких правовых проверок, она как минимум обязывает государства-участники установить соответствующую формальную процедуру. По определению МККК это «означает, что будет существовать постоянно действующий механизм, в любое время готовый для осуществления проверки новых видов оружия, когда оно изучается, разрабатывается, приобретается или принимается на вооружение»¹²³. Например, в отношении автономных систем оружия Директива 3000.09 Минобороны США требует установления высоких стандартов их проектирования и испытания, а также подготовки соответствующего личного состава; обеспечения проверки на уровне высшего командования до разработки и развертывания в полевых условиях; а также запрета на использование такого оружия любым способом, который не предусматривался при проектировании, испытании, сертификации оружия, подготовке операторов и определении военной доктрины¹²⁴. Несмотря на то что США не являются участником ДП I, изложенные в Директиве 3000.09 стандарты могут служить полезным примером практической организации формальных процедур испытания в соответствии со статьей 36. Это особенно актуально в свете того, что лишь малая толика государств располагает систематическими подходами к правовой проверке новых вооружений¹²⁵, а действующих механизмов проверки на наличие человеческих предубеждений, которые могут повлиять на работу моделей МО в системах вооружений, нет ни у одной страны мира (по крайней мере, по общедоступной информации).

122 ICRC, *Autonomy, Artificial Intelligence and Robotics: Technical Aspects of Human Control*, Geneva, August 2019, p. 3, доступно по адресу: <https://bit.ly/3a8787w>.

123 Руководство МККК по проверке новых видов оружия (примечание 105 выше). С. 20.

124 DoD, *Autonomy in Weapon Systems: Directive 3000.09*, 21 November 2012, доступно по адресу: <https://bit.ly/2XVRDtW>; DoD, *Law of War Manual*, Washington, DC, December 2016 (DoD Law of War Manual), § 6.5.9.4, доступно по адресу: <https://bit.ly/3sFrrBJ>.

125 Руководство МККК по проверке новых видов оружия (примечание 105 выше), с. 3 и 5; James D. Fry, "Contextualized Legal Reviews for the Methods and Means of Warfare: Cave Combat and International Humanitarian Law", *Columbia Journal of Transnational Law*, Vol. 44, No. 2, 2006, pp. 453, 473–479.

Принцип индивидуальной ответственности

Другая проблема связана с подотчетностью и принципом индивидуальной ответственности. МГП устанавливает набор принципов и норм, которые регулируют вопросы методов и средств ведения войны, в том числе личную уголовную ответственность за военные преступления¹²⁶. Как отмечает МККК, «нормы международного гуманитарного права адресованы людям. Именно люди соблюдают и имплементируют эти нормы, именно люди будут привлекаться к ответственности за их нарушение»¹²⁷.

Внедрение моделей МО в механизмы оценки боевых целей, очевидно, может иметь далеко идущие последствия и в конечном счете повлияет на парадигму обеспечения подотчетности. Другие значительные последствия может вызывать применение МО для поддержки принятия решений в контексте автономного оружия. Такое применение связано с дополнительными проблемами с точки зрения права, так как при применении норм МГП ответственность за вынесение определенных суждений несут люди. В этой связи возникают потенциальные проблемы подотчетности тех видов оружия или систем, принимающих решения, которые полностью полагаются на модели МО и не предполагают человеческого вмешательства. К сожалению, единого мнения о том, как следует решать такие проблемы, не существует.

МККК и ряд стран, например, пришли к выводу, что контроль со стороны человека необходим в любом случае, — чтобы полностью предотвратить появление пробелов в подотчетности: «...только на комбатантов возложена обязанность принимать решения в соответствии с нормами международного гуманитарного права, регулирующими ведение военных действий, и эту ответственность невозможно переложить на машину, компьютерную программу или алгоритм»¹²⁸. К такому же заключению пришли и все высокие договаривающиеся стороны Конвенции о запрещении или ограничении применения конкретных видов обычного оружия¹²⁹. Однако ряд государств и экспертов, напротив, придерживаются мнения о том, что вопросы подотчетности не являются особенно важными и что при разрывании автономного оружия достаточно некоторого «приемлемого уровня человеческого суждения», не уточняя при этом, относительно чего должно быть сформулировано это суждение¹³⁰. Вероятно, такие государства

126 Обычное МГП (примечание 112 выше), норма 151.

127 МККК (примечание 5 выше), р. 7. См. также: DoD Law of War Manual (примечание 124 выше), §6.5.9.3

128 МККК (примечание 5 выше). См. также: Eric Talbot Jensen, “The (Erroneous) Requirement for Human Judgment (and Error) in the Law of Armed Conflict”, *International Law Studies*, Vol. 96, No. 1, 2020, pp. 37–42 (где обобщаются позиции нескольких государств по причинам необходимости контроля со стороны человека).

129 Доклад сессии 2019 года Группы правительственных экспертов по вопросам, касающимся новых технологий в сфере создания смертоносных автономных систем вооружений, док. ООН CCW/GGE.1/2019/3, 25 сентября 2019 г., приложение IV, п. (b).

130 E. T. Jensen (примечание 128 выше), pp. 42–44.

чрезмерно полагаются на содержание статьи 36, которая устанавливает, что люди не освобождаются от ответственности за результат работы моделей МО, приведший к нарушению норм МГП, а также фактически любой нормы международного права. Как подчеркивается в комментарии к статье 36, если предусмотренные ею меры предосторожности не были приняты, государство в *любом* случае несет ответственность за противоправное причинение любого вызванного этим ущерба¹³¹.

Предполагается, однако, что такое восприятие слишком сильно опирается на процесс проверки вооружений. Безусловно, статья 36 регламентирует меры предосторожности в период до применения новых видов оружия, однако в ситуации реального применения новых методов или средств ведения войны в ходе вооруженного конфликта применимость ее положений будет ограничена. Так, в ней не конкретизируется порядок проверки легитимности оружия, методов и средств ведения войны. Более того, она не предусматривает механизмов привлечения к ответственности за какие-либо непредвиденные последствия применения нового оружия. Эти недостатки особенно осложняют ситуацию, когда речь идет о МО. Системы МО «по определению непредсказуемы» в том смысле, что они постоянно обучаются и адаптируются к изменениям, исходя из обрабатываемых ими данных. Кроме того, «машины не имеют человеческого представления о природе или концепции» наблюдаемых ими объектов¹³². Ввиду двух данных обстоятельств модель МО, полностью отвечающая всем требованиям в области МГП на этапе проверки оружия, при применении в полевых условиях может «ошибиться» или «сработать неправильно» так, что это приведет к нанесению ущерба гражданским лицам¹³³. В такой ситуации может возникнуть вакуум ответственности, так как, согласно МГП, для наступления личной ответственности субъект должен действовать как минимум «с осознанием того, что обстоятельство существует или что последствие наступит при обычном ходе событий»¹³⁴.

Аргументом против такой позиции является утверждение о том, что включение человеческих предубеждений в модели МО неизбежно (мы выяснили это в предыдущем разделе), в силу чего такие предубеждения и их последствия являются полностью предвидимыми. Тем не менее многие судьи международных судов не стремятся выносить обвинительные приговоры на основании столь общих и неконкретных аргументов и при отсутствии доказательств того, что подсудимый был осведомлен о конкретных предшествующих обстоятельствах, в которых наступили соответствующие последствия¹³⁵. К примеру, при рассмотрении дела Бембы большая часть

131 ICRC Commentary on APs (примечание 106 выше), para. 1466 (курсив наш).

132 ICRC (примечание 102 выше), p. 13. См. также: S.-S. Hua (примечание 98 выше), pp. 128–129.

133 S.-S. Hua (примечание 98 выше), pp. 128–129.

134 Римский статут (примечание 19 выше), ст. 30.

135 См., например: ICTY, *Prosecutor v. Milan Milutinović et al.*, Case No. IT-05-87-T, Judgment (Trial Chamber), 26 February 2009, para. 933.

членов Апелляционной палаты оправдала обвиняемого¹³⁶. В особом мнении двое судей, принадлежащих к упомянутому большинству, определили, что информация о преступлениях, совершенных подчиненными обвиняемого, в тот момент не была «достаточно конкретной». По их утверждению, «для обвинения командующего недостаточно знать, что он был осведомлен об общем риске совершения военными, находящимися в его подчинении, каких-либо преступлений», так как «такой риск существует при проведении любой военной операции»¹³⁷. Очевидно, что ответственные за развертывание модели МО, которая «случайно» выбирает гражданских лиц в качестве боевых целей из-за содержащихся в ней предубеждений, могут привести аналогичный аргумент и указать на отсутствие каких-либо конкретных «ошибок» на этапе испытаний как обоснование того, что они не были в правовом смысле осведомлены о преступлении. Изложенные соображения позволяют с легкостью опровергнуть любое предположение о том, что проведения процедуры правовой проверки вооружений достаточно для того, чтобы юридически доказать факт обеспечения правовой подотчетности.

В целом, с учетом возможных последствий наличия предубеждений, с точки зрения основополагающих норм МГП совершенно необходимо, чтобы как государства, так и международные организации активно работали над повышением эффективности процесса выработки политики в отношении всех аспектов этой проблемы. Выработывая соответствующую политику, следует сосредоточиться на создании механизмов и методик обеспечения подотчетности в тех случаях, когда применение моделей МО ведет к нарушению норм МГП.

Правовые последствия предубеждений с точки зрения МУП

Модели МО и инструменты анализа больших данных, затронутые любым из упомянутых выше типов предубеждений, могут оказывать реальное воздействие на расследование международных преступлений и ход судебных разбирательств. Из-за них может формироваться недостоверная картина обстоятельств на месте событий, могут закрепляться отрицательные стереотипы или другие расовые или гендерные предубеждения либо даже ложная оправдывающая информация. Таким же образом модели МО, содержащие предубеждения, могут закреплять предубеждения и создавать условия для их самовоспроизводства. Посредством такого комплексного влияния вышеописанные предубеждения могут приводить к серьезным правовым последствиям, воздействуя на обязанности Прокурора и права обвиняемых

136 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo*, Case No. ICC-01/05-01/08, Judgment on the appeal of Mr Jean-Pierre Bemba Gombo against Trial Chamber III's "Judgment Pursuant to Article 74 of the Statute" (Appeals Chamber), 8 June 2018.

137 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo*, Case No. ICC-01/05-01/08, Separate Opinion of Judge Christine Van den Wyngaert and Judge Howard Morrison (Appeals Chamber), 8 June 2018, para. 44.

и пострадавших лиц. В следующем разделе с опорой на Римский статут рассматриваются некоторые из таких последствий; основное внимание уделено потенциальному воздействию предубеждений на исполнение Прокурором своих обязанностей по установлению истины и изучению доказательств невиновности, на соображения о применимости тех или иных доказательств, а также на обязательства Прокурора о раскрытии.

Воздействие на обязанность Прокурора по установлению истины

Одна из особенностей Римского статута заключается в том, что он возлагает на Прокурора обязанность, во исполнение которой последний «для установления истины проводит расследование с тем, чтобы охватить все факты и доказательства, относящиеся к оценке того, наступает ли уголовная ответственность в соответствии с настоящим Статутом»¹³⁸. Для этого Прокурор «в равной мере расследует обстоятельства, свидетельствующие как о виновности, так и о невиновности»¹³⁹. Еще одно проявление этой идеологии можно увидеть в статье 81(1)(b) Римского статута, которая позволяет Прокурору обжаловать приговор от лица обвиняемого. Как отмечается в одном из комментариев, в силу подобной природы Римского статута Прокурор превращается из «заинтересованной стороны судебного процесса в субъект правосудия»¹⁴⁰.

Принципиально важно, что эти обязательства создают юридическое требование, предусматривающее, что проводимое Прокурором расследование должно быть достаточно глубоким и нейтральным, чтобы «установить истину» и обеспечить сбор информации, которая может опровергнуть вину. Относительно событий в Афганистане члены Апелляционной палаты отметили, что

для получения полной картины соответствующих деяний, обоснования того, что они с точки зрения права могут быть отнесены к конкретным типам преступлений, входящим в компетенцию Суда, и установления ответственности ряда потенциально вовлеченных в их совершение субъектов Прокурор должен расследовать ситуацию в целом, а не только подтверждающие вину обстоятельства¹⁴¹.

138 Римский статут (примечание 19 выше), ст. 54(1).

139 Там же.

140 Claus Kress, “The Procedural Law of the International Criminal Court in Outline: Anatomy of a Unique Compromise”, *Journal of International Criminal Justice*, Vol. 1, No. 3, 2003, p. 608.

141 ICC, *Situation in the Islamic Republic of Afghanistan*, Case No. ICC-02/17, Judgment on the Appeal against the Decision on the Authorisation of an Investigation into the Situation in the Islamic Republic of Afghanistan (Appeals Chamber), 5 March 2020, para. 60. См. также: ICC, *Prosecutor v. Thomas Lubanga Dyilo*, Case No. ICC-01/04-01/06, Judgment on the Prosecutor’s Appeal against the Decision of Pre-Trial Chamber I Entitled “Decision Establishing General Principles Governing Applications to Restrict Disclosure Pursuant to Rule 81(2) and (4) of the Rules of Procedure and Evidence” (Appeals Chamber), 12 October 2006, para. 52.

Подобное обязательство также отражено в программных документах и кодексах КП. Например, согласно Программному документу КП по преступлениям на сексуальной и гендерной почве Канцелярия «должна расследовать обстоятельства, свидетельствующие как о виновности, так и о невиновности, относительно преступлений на сексуальной и гендерной почве справедливо и беспристрастно для установления истины»¹⁴². В Кодексе поведения КП аналогичным образом указано, что во исполнение этих обязательств члены КП должны «рассмотреть все относящиеся к делу обстоятельства при оценке доказательств, независимо от того, подкрепляют ли они позицию обвинения или опровергают ее»¹⁴³.

Применительно к расследованию на основании больших данных такие обязательства, вероятно, требуют от Прокурора принятия мер, препятствующих воздействию предубеждений на любые модели МО или инструменты анализа больших данных. Прокуроров, занимающих должности в учреждениях МУП, традиционно критикуют за неумение работать с доказательствами невиновности¹⁴⁴. Если прокуроры будут полагаться на модели МО и большие данные, не принимая достаточных мер предосторожности против предубеждений, которые могут предотвратить получение или раскрытие доказательств невиновности, подобная критика станет еще более острой. Как показано выше, если модель МО основывается на данных, в которых отражены те или иные предубеждения, принимаемые исходя из таких данных решения будут систематически ущемлять интересы тех групп, которые чрезмерно или недостаточно представлены в соответствующем наборе данных. Аналогичным образом, если в методологии обследования источников больших данных не учтены потенциальные предубеждения средств расследования и следователей, в ходе расследования с большой долей вероятности будет упущена оправдывающая информация или информация о других преступлениях, например о ПСГП, которые, как правило, недостаточно представлены в крупных наборах данных. Подобные последствия явным образом сказываются на способности Прокурора исполнять свои обязательства в соответствии со статьей 54(1) Римского статута.

Риск не заметить или упустить доказательства невиновности сопряжен с особенно тяжелыми последствиями. Как было отмечено при оценке моделей МО, применяемых местными правоохранительными органами в США, «при том что органы прокуратуры, как и весь остальной мир, начинают ориентироваться на данные в своей работе, [системы сбора больших данных] не предусматривают возможности выявления доказательств неви-

142 OTP Policy Paper on SGBC (примечание 80 выше), para. 48.

143 ICC, OTP, *Code of Conduct for the Office of the Prosecutor*, 5 September 2013, para. 49(b), доступно по адресу: <https://bit.ly/3itiSoU>.

144 См., например: ICTY, *Prosecutor v. Radovan Karadžić*, Case No. IT-95-5/18-T, Decision on Accused's Ninety-Fourth Disclosure Violation Motion (Trial Chamber), 13 October 2014; ICC, *Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaïssona*, Case No. ICC-01/14-01/18, Decision on the Yekatom Defence Request Concerning Disclosure Violation (Trial Chamber V), 18 January 2021.

новности или порочащих свидетеля доказательств для защиты»¹⁴⁵. Такая ситуация обусловлена тем, что аналитические инструменты, используемые при расследованиях, обычно настроены на доказывание вины. К примеру, большинство криминологов-аналитиков сегодня используют системы управления данными для визуализации сетей, проведения анализа социальных сетей и отслеживания геопространственных или временных отношений, чтобы более эффективно обнаруживать скрытые связи и закономерности в данных (помимо прочего). Аналитики Международного уголовного трибунала по бывшей Югославии (МТБЮ) самым активным образом использовали ПО Analyst's Notebook, ZyFIND и CaseMap, что позволило им структурировать и категоризировать информацию с помощью аналитических примечаний или тегов (ярлыков) в единой базе данных, а затем разными способами отфильтровывать и по-новому структурировать ту же информацию. За счет этого процесс анализа был значительно упрощен¹⁴⁶. Похожая ситуация наблюдается в МУС: аналитики КП используют базу данных для анализа фактов в качестве вспомогательного инструмента при расследованиях, чтобы сопоставлять и объединять все источники доказательств о релевантных группах, местах, лицах и других субъектах и объектах¹⁴⁷.

Ни один из этих инструментов не сможет сравниться по сложности с моделями МО или аналитическими базами данных, применяемыми в некоторых странах или частных корпорациях, но тем не менее они помогают сформировать определенное представление о потенциальных проблемах. В частности, такие системы обычно создаются для выявления отношений, указывающих на вину того или иного человека. Так, криминологи-аналитики, в том числе и работающие в МУС, сегодня гораздо чаще сопоставляют коммуникационные данные из журналов вызовов, писем электронной почты, переписок в социальных сетях и других средств общения, чтобы выявлять социальные сетевые формирования, информацию о которых, в свою очередь, можно использовать для логического выяснения организационных отношений, в том числе иерархии¹⁴⁸. Подобный анализ неочевиден при определении связей между военными или политическими руководителями и теми, кто выполняет их указания на местах. Но опять же, эти аналитические инструменты в основном предназначены для обнаружения преступных сетей и определения роли отдельных лиц в этих сетях, а не для опровержения факта их существования. В этой связи их использование может дать почву для якобы правдоподобного заявления о том, что

145 Andrew G. Ferguson, "Big Data Prosecution and *Brady*", *UCLA Law Review*, Vol. 67, No. 1, 2020, p. 184.

146 Richard A. Wilson and Matthew Gillett, *The Hartford Guidelines on Speech Crimes in International Criminal Law*, 2018, para. 265, доступно по адресу: <https://bit.ly/2M0To6e>.

147 ICC, *Proposed Programme Budget for 2018 of the International Criminal Court*, ICC-ASP/16/10, 11 September 2017, para. 330, доступно по адресу: <https://bit.ly/3itlDXi>.

148 Nema Milaninia, "Using Mobile Phone Data to Investigate Mass Atrocities and the Human Rights Considerations", *UCLA Journal of International Law and Foreign Affairs*, Vol. 24, No. 2, 2020, pp. 283–286.

Прокурор не «в равной мере расследует обстоятельства, свидетельствующие как о виновности, так и о невинности»¹⁴⁹ и действует в нарушение статьи 54(1) Римского статута.

Огромный объем имеющейся информации также увеличивает вероятность того, что следователи или прокуроры не заметят или упустят доказательства невинности ввиду скрытых предубеждений, а это, в свою очередь, может сказаться на результатах работы любых применяемых моделей МО или инструментов анализа больших данных. Сегодня прокуроры и следователи работают с беспрецедентно большими объемами информации. Это особенно верно в отношении расследования международных преступлений, которые часто состоят из множества географически и хронологически неоднородных инцидентов с участием тысяч субъектов¹⁵⁰. Ситуация усугубляется еще и тем, что на следователей и прокуроров оказывается значительное давление с тем, чтобы они успешно выполняли свои обязанности в рамках обвинения, а не пытались установить истину¹⁵¹.

Неудивительно, что в таких условиях скрытые предубеждения только обостряют существующие проблемы. Даже если следователи и прокуроры обнаружили доказательства, которые непосредственно указывают на то, что обвиняемый находился на месте преступления, видеозаписи и фотографии, распространяемые с различных аккаунтов в социальных сетях, могут поставить правдоподобность таких доводов под вопрос. Доказательства, указывающие на виновность, например сообщения с угрозами, могут смешиваться с доказательствами невинности, такими как фотографии обвиняемого в удаленных от места преступления местах с временными метками. Учитывая то, как много времени и усилий требуется для извлечения и тщательной проверки такого объема информации, основная часть времени в рамках расследования уделяется сбору именно доказательств виновности, а не наоборот. Наконец, по мере роста объема данных любому прокурору или следователю становится сложнее глубоко понять связи между различными доказательствами, в том числе в тех случаях, когда некоторые доказательства указывают на недостоверность показаний свидетеля или противоречат его заявлениям. Чем больше и сложнее становятся данные и отношения между свидетелями, местами и группами, тем тяжелее определить, является ли та или иная информация существенной для привлечения обвиняемого к ответственности.

Наконец, на предубеждения в моделях МО и инструментах анализа больших данных потенциально может влиять недостаточный учет пре-

149 Elizabeth E. Joh, "The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing", *Harvard Law and Policy Review*, Vol. 10, No. 1, 2016, p. 25; Jennifer A. Johnson, John David Reitzel, Bryan F. Norwood, David M. McCoy, D. Brian Cummings and Renee R. Tate, "Social Network Analysis: A Systematic Approach for Investigating", *FBI Law Enforcement Bulletin*, 5 March 2013, доступно по адресу: <https://bit.ly/35SfIH6>.

150 IER Report (примечание 92 выше), para. 479.

151 ICC, OTP, *Regulations of the Office of the Prosecutor*, ICC-BD/05-01-09, 23 April 2009, Regulation 8; ICC, *Proposed Programme Budget for 2020 of the International Criminal Court*, ICC-ASP/18/10, 25 July 2019, para. 278, доступно по адресу: <https://bit.ly/39HayOs>.

ступлений или склонность общества к их стигматизации. Расследование убийств, совершаемых в дневное время и запечатленных на камеры мобильных телефонов, для следователей — линия наименьшего сопротивления, идя по которой, они никогда не узнают о случаях сексуального насилия, совершаемого вдали от посторонних глаз¹⁵². Предубеждения в отчетности и ошибки отбора серьезно снижают видимую значимость и масштаб недостаточно учитываемых преступлений, в особенности ПСП. Модели, основанные на таких данных, могут нести в себе упомянутые предубеждения и ограничивать возможности Прокурора по выявлению и расследованию малоосвещаемых преступлений, а также по «установлению истины» в их отношении. Так, например, исследователи медицинского факультета Гарвардского университета создали модель выявления факторов риска, которая позволит предотвращать случаи сексуального насилия в отношении женщин-военнослужащих в сухопутных войсках США. При ее разработке они опирались на административные доклады и обследования для выявления случаев сексуального насилия, сделав при этом поправку в наборе данных на те случаи, о которых не заявляют жертвы и которые не удастся выявить в рамках обследований. Выяснилось, что «в государственные органы было заявлено не более чем о 29,3% всех случаев сексуального насилия, совершенного в отношении таких женщин, не более чем о 34,2% случаев женщины сообщили сами в рамках обследования, и не более 46,5% случаев нашли отражение либо в заявлениях в госорганы, либо в обследовании»¹⁵³. Без аналогичных корректировок в наборах данных по недостаточно учитываемым международным преступлениям любые результаты анализа будут неизбежно искажены в пользу более освещаемых событий, что препятствует исполнению обязательства Прокурора по установлению истины в соответствии со статьей 54 Римского статута.

Соображения относительно допустимости доказательств

Еще одним аспектом, на который предубеждения в моделях МО и больших данных могут оказывать явное воздействие, является принятие доказательств к рассмотрению. Согласно статье 69(4) Римского статута МУС может выносить решения об относимости или допустимости любых доказательств. В этой связи Суд принимает во внимание неисчерпывающий перечень факторов, в том числе «их силу, а также вред, который такие доказательства могут причинить проведению справедливого судебного разбирательства или справедливой оценке показаний свидетеля».

Это положение оставляет судьям широкие возможности по оценке предубеждений, которые могли повлиять на сбор или анализ информации, при определении их воздействия на показания свидетеля или даже

152 N. Milaninia (примечание 148 выше), p. 297.

153 Amy E. Street *et al.*, “Developing a Risk Model to Target High-risk Preventive Interventions for Sexual Assault Victimization among Female U.S. Army Soldiers”, *Clinical Psychological Science*, Vol. 4, No. 6, 2016.

ход судебного процесса. Кроме того, оно позволяет судьям оценивать и учитывать такие предубеждения при определении того, какой вес имеет то или иное доказательство или является ли оно в принципе допустимым. Например, судьи могут заключить, что неспособность следователей КП учесть предубеждения при сборе доказательств в Facebook или Twitter могла препятствовать изучению или формированию ими сведений, которые потенциально опровергают показания того или иного свидетеля.

Чаще всего предубеждения влияют на определение веса экспертных заключений относительно результатов работы моделей МО или результатов анализа больших наборов данных, например относительно закономерностей в преступлениях. Вес экспертного заключения зависит исключительно от надежности информации, на которую эксперт опирается при подготовке выводов. Если эксперт сделал(-а) выводы, исходя из больших наборов данных или результатов работы алгоритма, который исключил или не учел релевантные доказательства, экспертное заключение, основанное на таком анализе, будет носить спорный характер. Подобные изыскания не новы — на национальном уровне, например, суды уже осведомлены о проблемах, связанных с методами сбора данных, и проводят оценку того, не привели ли эти методы к формированию искаженной выборки, которая снизит значимость полученных данных применительно к рассматриваемому вопросу. Это особенно актуально в отношении мнений и выводов, формируемых с использованием баз доказательных данных, таких как Analyst's Notebook. Ныне покойная Судья Верховного суда США Рут Бейдер Гинзбург признавала, что «риск ошибки, проистекающий из... [использования] баз данных, отнюдь не мал», отмечая проблемы, возникающие в работе Национального информационно-криминологического центра, а также при анализе перечней лиц, подозреваемых в террористической деятельности, и баз данных государственной службы занятости¹⁵⁴.

Наконец, присутствие предубеждений в моделях МО также может повлиять на выбор доказательного аппарата судьями по конкретному делу, если они осведомлены о таких предубеждениях и понимают их природу. Римский статут оставляет достаточно большой простор для выбора методик оценки доказательств. Апелляционная палата МУС подтвердила, что согласно статье 69(4) Статута Судебная палата вправе определять, будет ли решение о допустимости того или иного доказательства принято сразу по его представлении в ходе судебного заседания (так называемый метод «допущения») или в конце рассмотрения дела, когда все доказательства были представлены и заслушаны (так называемый метод «подачи»)¹⁵⁵. В последнее время ряд Судебных палат все чаще придерживается

154 US Supreme Court, *Herring v. United States*, 555 U.S. 135, Justice Ginsburg Dissenting, 14 January 2009, p. 155.

155 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Judgment on the Appeals of Mr Jean-Pierre Bemba Gombo, Mr Aimé Kilolo Musamba, Mr Jean-Jacques Mangenda Kabongo, Mr Fidèle Babala Wandu and Mr Narcisse Arido against the Decision of Trial Chamber VII Entitled “Judgment Pursuant to Article 74 of the Statute” (Appeals Chamber), 8 March 2018, paras 576–601;

второго подхода, согласно которому Судебная палата откладывает вынесенные любых решений о релевантности, доказательной силе и потенциальной предвзятости любых доказательств до конца судебного разбирательства и начала процесса принятия окончательного решения согласно статье 74(2) Римского статута. В этот момент Судебная палата изучает каждое доказательство, «поданное» в течение процесса, с точки зрения всех стандартных критериев доказательности¹⁵⁶. В практическом смысле выбор такого подхода Судебной палатой позволяет «подавать» любые доказательства в ходе разбирательства, даже если имеются основания полагать, что эта информация неподлинна или недостоверна.

В обстоятельствах, когда Прокурор планирует опираться на большие данные или на модели МО и инструменты анализа, основанные на источниках больших данных, представляется, что Судебная палата должна использовать метод «подачи», а не определять допустимость доказательств в ходе судебного процесса. Такое соображение обусловлено тем, что из-за разнообразия потенциальных индикаторов достоверности, применение которых может повлиять на вес или релевантность свидетельских показаний или других доказательств с точки зрения судей, последние, скорее всего, предпочтут обзреть весь объем доступной информации, не отклоняя поспешно те или иные отдельные сведения. Такая осмотрительность будет разумна при изучении потенциального воздействия предубеждений на процесс расследования, и, в частности, на возможное исключение оправдывающих материалов.

Обязательства о раскрытии

Последней сферой, в которой предубеждения в моделях МО и инструментах анализа больших данных могут вызывать правовые последствия с точки зрения МУП, являются вопросы раскрытия. В Римском статуте подробно определены обязанности по раскрытию. Прокурор обязан раскрывать защите «так скоро, насколько это возможно» и на регулярной основе имеющиеся в его распоряжении или под его контролем доказательства,

ICC, *Prosecutor v. Jean-Pierre Bemba Gombo*, Case No. ICC-01/05-01/08, Judgment on the Appeals of Mr Jean-Pierre Bemba Gombo and the Prosecutor against the Decision of Trial Chamber III Entitled “Decision on the Admission into Evidence of Materials Contained in the Prosecution’s List of Evidence” (Appeals Chamber), 3 May 2011, para. 37.

156 ICC, *Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaijssona*, Case No. ICC-01/14-01/18, Initial Directions on the Conduct of the Proceedings (Trial Chamber V), 26 August 2020, paras 52–53; ICC, *Prosecutor v. Al Hassan Ag Abdoul Aziz Ag Mohamed Ag Mahmoud*, Case No. ICC-01/12-01/18, Annex A to the Decision on the Conduct of Proceedings (Trial Chamber X), 6 May 2020, paras 30–31; ICC, *Prosecutor v. Dominic Ongwen*, Case No. ICC-02/04-01/15, Initial Directions on the Conduct of the Proceedings (Trial Chamber IX), 13 July 2016, paras 24–25; ICC, *Prosecutor v. Laurent Gbagbo and Charles Blé Goudé*, Case No. ICC-02/11-01/15, Decision on the Submission and Admission of Evidence (Trial Chamber I), 29 January 2016; ICC, *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Decision on Prosecution Requests for Admission of Documentary Evidence (ICC-01/05-01/13-1013-Red, ICC-01/05-01/13-1113-Red, ICC-01/05-01/13-1170-Conf) (Trial Chamber VII), 24 September 2015, paras 10–13.

которые, по его мнению, свидетельствуют или должны свидетельствовать о невиновности обвиняемого, или о наличии обстоятельств, смягчающих вину, или способны повлиять на достоверность доказательств обвинения (статья 67(2) Римского статута), или имеют существенное значение для подготовки защиты (правило 77 Правил процедуры и доказывания МУС). Существует три возможных пути воздействия предубеждений в моделях МО и инструментах анализа больших данных на обязательства Прокурора о раскрытии.

Во-первых, методология, используемая обвинением для сбора информации из источников больших данных или полученная от любой модели МО, может подлежать раскрытию в соответствии с правилом 77 Правил процедуры и доказывания МУС. Данное обстоятельство, однако, подразумевает, что защита будет осведомлена о некоторых предубеждениях, которые могли воздействовать на ход проводимого обвинением анализа или сбора данных. Это объясняется тем, что правило 77 не требует раскрытия той информации, которая является существенной лишь потенциально. Правило не дает защите возможности начать спекулятивный подбор потенциально компрометирующей информации¹⁵⁷; скорее, оно обязывает защиту привести достаточное подтверждение существенности¹⁵⁸. Для этого защита должна будет предоставить подробное и достаточно конкретное обоснование существенности запрашиваемых материалов¹⁵⁹.

Если защите удастся добиться достаточной конкретности, информация о той модели МО или методологическом подходе, которыми пользовалась КП при оценке больших данных, вероятно, должна быть признана «существенной», хотя до настоящего времени такой правоприменительной практики в МУС зафиксировано не было. При этом в деле Бембы и др. Судебная палата заключила, что просьбы об оказании помощи (ПОП) — письма, направляемые в адрес государств, с запросом на получение доказательных материалов — были существенными для подготовки защиты и подлежали раскрытию в соответствии с правилом 77 Правил процедуры и доказывания, «будучи неразрывно связанными с допустимостью доказательств, на которые опиралось обвинение». В упомянутых просьбах об оказании помощи запрашивались записи перехваченных коммуникаций

157 См.: ICC, *Prosecutor v. Thomas Lubanga Dyilo*, Case No. ICC-01/04-01/06, Decision on the Defence Request for Unrestricted Access to the Entire File of the Situation in the Democratic Republic of the Congo (Pre-Trial Chamber I), 17 May 2006, pp. 2–3 (которым был отклонен запрос защиты на получение доступа к полному досье о ситуации в ДРК с указанием на ходатайство обвинения о том, что данный запрос является попыткой подобрать потенциально компрометирующие материалы и не имел под собой законных судебных оснований). См. также: ICTY, *Prosecutor v. Dragomir Milošević*, Case No. IT-98-29/1-A, Decision on Motion Seeking Disclosure of Rule 68 Material (Trial Chamber I), 7 September 2012, para. 5.

158 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo*, Case No. ICC-01/05-01/08, Decision on Defence Requests for Disclosure (Trial Chamber III), 2 July 2014, para. 29.

159 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Decision on Mangenda Defence Request for Cooperation (Trial Chamber VII), 14 August 2015, para. 11; ICC, *Prosecutor v. Saif Al-Islam Gaddafi and Abdullah Al-Senussi*, Case No. ICC-01/11-01/1, Corrigendum to Decision on the “Defence Request for an Order of Disclosure”, (Pre-Trial Chamber I), 1 August 2013, para. 40.

обвиняемого и соответствующие записи журналов вызовов¹⁶⁰. Палата пришла к выводу о том, что, так как защита намеревалась оспорить правомерность ПОП как несоразмерных или основанных на вводящей в заблуждение информации, «защите следовало обеспечить возможность проверить надежность процедур, использованных для сбора доказательств виновности»¹⁶¹. Такой аргумент соответствовал содержанию ранее вынесенного Палатой определения, где она заключала, что «материалы, позволяющие защите оценить легитимность доказательств, на которые в рамках процесса планирует полагаться обвинение, являются существенными для подготовки защиты»¹⁶².

Подобный аргумент, вероятно, может быть применен к процессу использования моделей МО следователями КП в рамках анализа или сбора больших данных. С учетом потенциального воздействия скрытых предубеждений, как подробно изложено выше, раскрытие такой информации или документов, в которых зафиксирован порядок выполнения таких процессов, можно рассматривать как «неразрывно связанное» с допустимостью доказательств, полученных в результате таких процессов. В таком случае они могут подлежать раскрытию в соответствии с правилом 77 аналогично ПОП в рамках дела Бембы и др. Однако это возможно только при условии, что раскрытие такой информации не запрещается любой другой нормой права, такой как, например, правило 81 Правил процедуры и доказывания, в котором от раскрытия явно защищаются доклады и другие внутренние документы КП (см. ниже).

Во-вторых, такие предубеждения могут оказать воздействие на исполнение Прокурором обязательства по раскрытию оправдывающих материалов в соответствии со статьей 67(2) Римского статута. В соответствии с этим положением Прокурор обязан раскрывать «имеющиеся в его распоряжении или под его контролем доказательства, которые, по его мнению, свидетельствуют или должны свидетельствовать о невиновности обвиняемого, или о наличии обстоятельств, смягчающих вину, или способны повлиять на достоверность доказательств обвинения». Правоприменительная практика МУС явно устанавливает, что обязательство Прокурора по раскрытию в соответствии со статьей 67(2) не требует от него целенаправленного сбора оправдывающих материалов, а лишь требует продемонстрировать имеющуюся у него информацию такого рода¹⁶³. Исходя из этого, хотя Прокурор и обязан самостоятельно расследовать

160 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Decision on Defence Requests for Prosecution Requests for Assistance, Domestic Records and Audio Recordings of Interviews (Trial Chamber VII), 10 September 2015, para. 13.

161 Ibid.

162 ICC, *Bemba Gombo et al.* (примечание 159 выше), para. 10.

163 ICC, *Prosecutor v. William Samoei Ruto, Henry Kiprono Kosgey and Joshua Arap Sang*, Case No. ICC-01/09-01/11, Decision on the Defence Requests in Relation to the Victims' Applications for Participation in the Present Case (Pre-Trial Chamber II), 8 July 2011, para. 9; ICC, *Prosecutor v. Callixte Mbarushimana*, Case No. ICC-01/04-01/10, Decision on Issues relating to Disclosure (Pre-Trial Chamber I), 30 March 2011, para. 15.

обстоятельства, свидетельствующие о невинности в соответствии со статьей 54(1), упущение таких обстоятельств не считается нарушением обязательства о раскрытии, так как в распоряжении Прокурора нет информации о таких обстоятельствах.

В правоприменительной практике весьма затруднительно найти четкое определение того, что означает нахождение информации «под контролем» Прокурора. Данная проблема может особенно часто проявляться тогда, когда идет речь об аналитической информации и доказательствах, полученных с помощью моделей МО или анализа больших данных. Например, с учетом того, что, используя один из имеющихся у них аналитических инструментов, аналитики КП способны подготавливать доклады, демонстрирующие отсутствие связи обвиняемого с непосредственными исполнителями преступления или его иерархическую отдаленность от них, можно сказать, что информация находится «под контролем» КП, даже если она ей не распоряжается. Однако, даже если вышеизложенное верно, такие доклады скорее всего будут защищены от раскрытия в соответствии с правилом 1 Правил процедуры и доказывания¹⁶⁴. Данное правило явно защищает от раскрытия «доклады, меморандумы и иные внутренние документы, подготовленные одной из сторон, ее помощниками или представителями в связи с расследованием или подготовкой дела». Тем не менее даже в таких обстоятельствах подобные доклады могут иметь значение для исполнения Прокурором обязательства, установленного статьей 54(1), как описано выше.

Наконец (в-третьих), такие предубеждения создают риск дальнейшего затруднения процесса раскрытия в рамках работы МУС, который и без того является достаточно запутанным. Ввиду ряда взаимоусиливающих проблем процесс раскрытия в МУС значительно затруднен, и это, в свою очередь, неоднократно становилось главным фактором увеличения продолжительности судебных разбирательств¹⁶⁵. Среди таких проблем можно назвать следующие.

- Расследования в сфере МУП обыкновенно имеют крупный масштаб и охватывают сотни, если не тысячи потенциальных преступных деяний, которые совершались на протяжении многих лет. Таким образом, доказательная база отличается чрезвычайно большим объемом и часто содержит сотни тысяч страниц различных документов (многие из которых написаны от руки), сотни часов аудио- и видеоматериалов и тысячи фотографий¹⁶⁶.

164 ICC, *Prosecutor v. Thomas Lubanga Dyilo*, Case No. ICC-01/04-01/06, Redacted Decision on the Prosecution's Disclosure Obligations Arising Out of an Issue Concerning Witness DRC-OTP-WWWW-0031 (Trial Chamber I), 20 January 2011, para. 16.

165 IER Report (примечание 92 выше), para. 481.

166 См., например: ICC, *Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaïssona*, Case No. ICC-01/14-01/18, Prosecution's Request to Vary the Decision on Disclosure and Related Matters (ICC-01/14-01/18-64-Red) (Pre-Trial Chamber II), 20 March 2019, para. 7.

- Некоторые палаты МУС, по-видимому, расширили объем обязательств Прокурора по раскрытию гораздо дальше, чем это предусмотрено установленным порядком. Так, правило 77 ограничивает обязательства Прокурора по раскрытию сведений, «которые имеют существенное значение для подготовки защиты или которые Прокурор намеревается использовать в качестве доказательства», — по трактовке некоторых судей это требование распространяется на любые сведения, которые «имеют достаточное значение» для защиты¹⁶⁷. Вследствие такой трактовки под угрозой раскрытия оказывается абсолютно вся или любая информация, что не соответствует четкой формулировке правила 77. И действительно, как минимум одна группы защиты пыталась (хотя и безуспешно) добиться применения правила 77 к любой информации, которая могла бы «помочь защите принять обоснованное решение о направлении запроса на допущение дополнительных доказательств, подтверждающих жалобу, а также подготовить этот потенциальный запрос»¹⁶⁸.
- Неоднозначность стандартов раскрытия и исключения информации палатами свидетельствует о том, что длительный процесс подготовки доказательств начинается только на поздних стадиях судебного процесса¹⁶⁹. Например, в деле Екатома и Нгайссоны Палата предварительного производства II разрешила обвинению не раскрывать деликатную информацию в доказательствах, указав в специальной таблице, содержатся ли в том или ином источнике материалы, свидетельствующие о вине или невиновности или подпадающие под правило 77¹⁷⁰. Годом позже при рассмотрении дела Али Кушайба та же судебная коллегия, исходя из этого подхода (и из обычной практики), потребовала от обвинения «пометить соответствующие разделы документов, заявлений и стенограмм как [свидетельствующие о вине], [свидетельствующие о невиновности] или [подпадающие под действие правила 77] или иным образом либо предоставить релевантную информацию путем указания номеров страниц и пунктов в специальных полях для метаданных»; это

167 ICC, *Prosecutor v. Abdallah Banda Abakaer Nourain and Saleh Mohammed Jerbo Jamus*, Case No. ICC-02/05-03/09 OA 4, Judgment on the Appeal of Mr Abdallah Banda Abakaer Nourain and Mr Saleh Mohammed Jerbo Jamus against the Decision of Trial Chamber IV of 23 January 2013 Entitled “Decision on the Defence’s Request for Disclosure of Documents in the Possession of the Office of the Prosecutor” (Appeals Chamber), 28 August 2013, para. 42.

168 ICC, *Prosecutor v. Jean-Pierre Bemba Gombo et al.*, Case No. ICC-01/05-01/13, Defence Request for Leave to Reply to the Prosecution’s Response to Bemba’s “Consolidated Request for Disclosure and Judicial Assistance”, ICC-01/05-01/13-2236-Conf, 6 October 2017, ICC-01/05-01/13-2236-Conf-Corr, 10 October 2017 (Appeals Chamber), 12 October 2017, para. 10.

169 IER Report (примечание 92 выше), para. 480 («Поступило ходатайство о том, чтобы на этапе проверки Прокурор не начинал редактирование и раскрытие прежде, чем Палата примет протокол редактирования»).

170 ICC, *Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaissona*, Case No. ICC-01/14-01/18, Public Redacted Version of “Decision on Disclosure and Related Matters” (Pre-Trial Chamber II), 23 January 2019. См. также: ICC, *Prosecutor v. Alfred Yekatom and Patrice-Edouard Ngaissona*, Case No. ICC-01/14-01/18, Prosecution’s Communication of the Disclosure of Evidence (Pre-Trial Chamber II), 31 July 2019.

чрезвычайно осложнило процесс раскрытия и потребовало большого физического труда¹⁷¹. Непредсказуемость требований к процедуре раскрытия, даже исходящих от одних и тех же судей, говорит о том, что обвинение может приступать к редактированию доказательств только тогда, когда введен в действие протокол раскрытия и исключения информации. В противном случае есть вероятность, что обвинению придется проделать всю эту работу заново, если Палата примет решение использовать новый подход — независимо от того, изменился ли состав коллегии судей.

- Технологические методы МУС относительно архаичны¹⁷², а это значит, что процесс обзора и маркировки источников, подлежащих раскрытию, проходит в основном в неавтоматическом режиме, несмотря на хроническую нехватку персонала и ограниченный бюджет Суда.
- Предубеждения в моделях МО и больших данных усугубляют все упомянутые проблемы. Обязательство «в равной мере расслед[овать] обстоятельства, свидетельствующие как о виновности, так и о невиновности», и устанавливая истину при должном исполнении означает, что добросовестное обвинение должно расширить свою работу по сбору данных так, чтобы учитывались любые предубеждения и информация, которая обычно не принимается во внимание. Увеличение объемов данных означает увеличение количества информации, которую надо рассмотреть, так как в силу обязательства по передаче доказательств, свидетельствующих о вине и виновности, прокурорам приходится проводить их поиск во всех имеющихся у них в распоряжении цифровых источниках, включая заключения судмедэкспертов. Это увеличивает и число источников данных. Как подчеркивается в заключении по итогам недавнего аудита МУС, опубликованном в Независимом экспертном обзоре, «разрешение вопросов раскрытия становится все более трудной задачей в свете широкого распространения материалов, имеющих отношение к событиям, которые рассматривает Суд в своих разбирательствах¹⁷³. В современном «мире больших данных» и ввиду того, что МУС планирует активнее полагаться на модели МО и аналитику больших данных, эта задача будет только усложняться. Работа над этими вопросами — создание более практичных и последовательных правил раскрытия и внедрение более совершенных технологий — сегодня становится все более необходимой.

171 См.: ICC, *Prosecutor v. Ali Muhammad Ali Abd-Al-Rahman ("Ali Kushayb")*, Case No. ICC-02/05-01/20, Second Order on Disclosure and Related Matters (Pre-Trial Chamber II), 2 October 2020, para. 24; ICC, *Prosecutor v. Ali Muhammad Ali Abd-Al-Rahman ("Ali Kushayb")*, Case No. ICC-02/05-01/20, Prosecution's Third Progress Report on the Evidence Review, Translation and Disclosure Process (Pre-Trial Chamber II), 9 October 2020, para. 25 (где отмечается, что «такой порядок приведет к значительному увеличению количества времени, которое требуется для первичного и вторичного обзора документов, потенциально подлежащих раскрытию, особенно объемных документов, таких как стенограммы допросов»).

172 IER Report (примечание 92 выше), paras 577–584.

173 Ibid., para. 479.

Заключение

Судья Верховного суда США Энтони Кеннеди однажды высказал следующую мысль: «Предубеждения легко приписать поведению других людей и сложно обнаружить в себе»¹⁷⁴. Неосознанные предубеждения проявляются в наших суждениях и отношении к другим. В организациях такие предубеждения могут закрепляться в форме систем, структур, политики и практических методов, создавая порочный круг, разорвать который чрезвычайно сложно. То же самое верно и для моделей МО и инструментов анализа больших данных, которые используют в своей работе и для формирования выводов информацию, в которой содержатся распространенные человеческие предубеждения.

Но даже с учетом этого активизация использования МО и инструментов анализа больших данных может приносить невероятную пользу. В сфере МГП модели МО способны сделать действия систем вооружений и прицеливания, а также военные решения более обоснованными и снизить потенциальное число гражданских жертв, к которым приводят такие действия и решения. Надлежащим образом разработанная модель МО, как было рассмотрено выше на примере МОСУ (CDEM), способна просчитать сопутствующий ущерб гражданским лицам и объектам куда более точно, нежели человеческий разум, снижая тем самым вероятность нанесения чрезмерного вреда гражданским лицам и их имуществу. Схожим образом в сфере МУП многие проблемы МУС — непоследовательность судебных решений, медленные темпы обзора доказательств, потенциально подлежащих раскрытию, неэффективность расследований, неудовлетворительные результаты расследований ввиду отсутствия доступа к информации в странах, где были совершены преступления, — могут быть решены с помощью моделей МО, и порой только с их помощью. И действительно, модели МО и инструменты анализа данных могут также служить «двойным ситом», которое отсеивает сами человеческие предубеждения, в противном случае влияющие на судопроизводство, и служит мерой предосторожности против таких предубеждений.

Вот почему было бы весьма мудрым решением озадачиться теми проблемами, которые могут возникнуть при большем распространении МО и аналитики больших данных. Самой важной из этих проблем является потенциальное воздействие на них распространенных человеческих предубеждений. Автор этой статьи стремится обеспечить точку опоры для разработки необходимых правил и практических методов, рассматривая вопрос присутствия предубеждений в моделях МО и инструментах анализа больших данных и подчеркивая потенциальные проблемы и их правовые последствия. С точки зрения МГП необходимы более подробные и обоснованные рекомендации со стороны международных организаций, таких как МККК, по поводу понимания, оценки и смягчения воздействия

174 US Supreme Court, *Williams v. Pennsylvania*, 136 S. Ct. 1899, 9 June 2016, p. 1905.

упомянутых предубеждений, а также относительно ограничений и методов применения моделей МО в системах вооружений и процессах принятия военных решений. Необходимы также разработка и внедрение соответствующих норм и рекомендаций в руководства по военному делу (схожих с Директивой 3000.09 Министерства обороны США, рассмотренной выше), соблюдение которых должно регулярно проверяться и отрабатываться. Аналогичным образом учреждениям МУП необходимо активнее действовать на упреждение, для того чтобы подготовиться к проблемам нейтрализации предубеждений в МО и больших данных. Это, в том числе, подразумевает потребность в непрерывном юридическом образовании и интеграции принципов и практических методов в руководство по деятельности КП.

Без таких упреждающих действий модели МО и инструменты анализа больших данных, какими бы привлекательными они ни были и сколь бы добросовестно ни планировали их использовать организации, закрепят в себе распространенные человеческие предубеждения и приведут к ущемлению интересов традиционно уязвимых и недостаточно представленных сообществ. В этом отношении их применение будет, очевидно, сопряжено с правовыми последствиями, в том числе с точки зрения основополагающих принципов МУП и МГП. Организации, которые обычно избегают контроля со стороны общественности, должны научиться быть открытыми к испытаниям, аудиту, критике и проверкам. Регулярная проверка таких систем, вероятно, является единственным способом сделать построенный на наших слабостях и пороках мир, в котором машины играют значительную роль, хотя бы в определенной степени справедливым и честным.