



Обработка данных при помощи искусственного интеллекта и риски дискриминации

Талапина Эльвира Владимировна

Институт государства и права РАН, 119019 Россия, Москва, ул. Знаменка, 10.
talapina@mail.ru, <https://orcid.org/0000-0003-3395-3126>

Аннотация

Дискриминация создает угрозу равенству как основному концепту правового государства. Эта проблема обретает новое звучание в связи с использованием искусственно-го интеллекта для принятия значимых юридических решений, поскольку искусственный интеллект способен принимать ошибочные решения в отношении конкретных лиц, в основе которых часто лежит дискриминация. Целью статьи стало изучение рисков дискриминации, чтобы учесть и избежать их в будущем правовом регулировании. Исследование основано на анализе доктринальных и нормативных источников разных стран, изучении имеющегося опыта использования искусственного интеллекта. Конкретным методом интеллектуального анализа данных является профилирование, которое не оставляет большого пространства для автономии и самоопределения личности. В связи с этим предлагается переоценить теорию информационного самоопределения, используя ее потенциал разделения ответственности между обладателем данных и обработчиком. Ввиду явных дискриминационных рисков профилирования некоторые операции уже запрещены (например, редлайнинг в США, генетическое профилирование в области страхования и занятости в ряде стран). Отдельно должны изучаться предиктивные способности искусственного интеллекта в отношении юридической оценки поведения конкретного человека. Опыт алгоритмического прогнозирования поведения человека, накопленный в рамках уголовного правосудия в США, свидетельствует о вероятностном характере такой оценки. Это способно нарушить права человека на справедливое судебное разбирательство и индивидуализацию наказания, если алгоритмическая оценка станет единственным основанием для вынесения судебного решения. В целом развитие приложений по решению рутинных правовых проблем, которые будут давать результаты на основании прошлых судебных решений, особо актуально в



странах общего права, где господствует прецедентное право . С учетом того, что Россия принадлежит к континентальной системе права перспективы использования американского опыта сомнительны . Рассмотрение конкретных видов недостатков, которые способны привести к дискриминационной обработке данных, позволило сделать выводы о контурах будущего законодательства в отношении деятельности искусственного интеллекта с учетом всех проанализированных рисков дискриминации .

Ключевые слова: искусственный интеллект, равенство, прозрачность, предиктивность, дискриминация, профилирование, алгоритм



Введение

В теории прав человека равенство и недискриминация являются основополагающими ценностями. В обоих случаях речь идет о равенстве прав с тем отличием, что право на недискриминацию имеет более узкое содержание и в этом смысле производно от права на равенство. Концепт равенства можно считать центральным принципом правового государства; в основе системы естественных прав человека лежит представление о том, что люди рождаются свободными и равными в своем достоинстве и правах [Лапаева В.В., 2008: 67]. Разумеется, в юридическом смысле речь идет именно о формальном равенстве как равенстве возможностей, обусловленных личными усилиями и личной волей субъектов права [Нерсисянц В.С., 2009:4].

Дискриминация посягает на равенство, представляя угрозу для правового государства. Дискриминация — это не просто различие в обращении. Не каждое различие незаконно и представляет собой дискриминацию. Также дискриминация — это не просто неравное обращение. Различие в обращении может быть незаконным, но не представлять собой дискриминации. Дискриминация имеет место в тех случаях, когда неблагоприятное различие в обращении является незаконным и основано на критерии, на основании которого закон запрещает проведение юридических различий [Mine M., 2003: 30].

В современный период дискриминация, осуществляемая людьми в отношении других людей, не искоренена, но уже выработаны определенные критерии ее обнаружения, доказывания и привлечения к ответственности виновных лиц. Однако с распространением автоматизированной обработки данных, в особенности на базе технологий больших данных, возникла другая проблема — дискриминации людей со стороны искусственного интеллекта. Вторгаясь в существующее правовое поле, искусственный интеллект затрагивает целый спектр имеющихся норм (защита данных, транспарентность, информационное самоопределение), а также создает новые риски принятия ошибочных решений в отношении конкретных лиц, в основе которых часто лежит дискриминация. Во всем мире сейчас ведется широкая научная дискуссия, в результате которой предстоит досконально изучить эти риски с тем, чтобы предложить недискриминационные механизмы обработки



данных искусственным интеллектом. Именно с учетом результатов такой дискуссии, на наш взгляд, следует развивать законодательство.

1. Нормативный контекст преодоления дискриминации как посягательства на равенство

Не будет преувеличением утверждение, что проблемам дискриминации посвящен целый массив нормативных актов, в особенности, на международном уровне. Подходы к дискриминации могут базироваться на общих параметрах (что характерно для более ранних актов общей направленности), так и на более специфических характеристиках, некоторые из которых возникли только в последние годы (например, генетическая дискриминация). Рассмотрим некоторые из них.

Статья 2 Всеобщей декларации прав человека (1948) устанавливает общий принцип недискриминации: «Каждый человек должен обладать всеми правами и всеми свободами, провозглашенными настоящей Декларацией, без какого бы то ни было различия, как-то: в отношении расы, цвета кожи, пола, языка, религии, политических или иных убеждений, национального или социального происхождения, имущественного, сословного или иного положения. Кроме того, не должно проводиться никакого различия на основе политического, правового или международного статуса страны или территории, к которой человек принадлежит, независимо от того, является ли эта территория независимой, подопечной, самоуправляющейся, или как-либо иначе ограниченной в своем суверенитете». Всеобщая декларация прав человека, провозглашая равенство перед законом и право на равную защиту закона, запрещает в связи с этим какую бы то ни было дискриминацию и подстрекательства к ней (ст. 7). Другой вид дискриминации запрещен в связи с оплатой труда, которая должна быть равной за равный труд (ст. 23). Таким образом, равенство образует и общий принцип, и самостоятельное право человека, а также служит базой для образования другого права — права на недискриминацию.

Контекст дискриминации в области занятости получил широкое развитие в актах Международной организации труда, основополагающим из которых является Конвенция о



дискриминации в области труда и занятий (1958) (Конвенция № 111). В ее трактовке термин «дискриминация» включает: а) всякое различие, недопущение или предпочтение, проводимое по признаку расы, цвета кожи, пола, религии, политических убеждений, национального происхождения или социальной принадлежности, приводящее к уничтожению или нарушению равенства возможностей или обращения в области труда и занятий; б) всякое другое различие, недопущение или предпочтение, приводящее к уничтожению или нарушению равенства возможностей или обращения в области труда и занятий, определяемое соответствующим членом по консультации с представительными организациями предпринимателей и трудящихся, где таковые существуют, и с другими соответствующими органами.

Расовая и гендерная дискриминация также получили разнообразную правовую поддержку — Конвенция ООН о ликвидации всех форм дискриминации в отношении женщин (1979), Международная конвенция о ликвидации всех форм расовой дискриминации (1965). Активное проникновение новых технологий в общественные отношения «освежило» тематику дискриминации, повлияв на формирование нового направления — генетической дискриминации. Назовем Международную декларацию ЮНЕСКО о генетических данных человека (16 октября 2003 г.), в которой закреплено стремление к тому, чтобы генетические данные человека и протеомные данные человека не использовались в целях, которые носят дискриминационный характер в силу того, что они нарушают или приводят к нарушению прав человека, основных свобод или человеческого достоинства отдельного лица, и чтобы они не использовались в целях, ведущих к стигматизации того или иного лица, семьи, группы или общины (ст. 7 (а)).

Стандарты международных актов повторяются или конкретизируются в региональных документах, целью которых также является обеспечение защиты от дискриминации. К их числу относится Американская конвенция о правах человека (1969), которая защищает от дискриминации по признаку расы, цвета кожи, пола, языка, религии, политических или иных убеждений, национального или социального происхождения, экономического положения, рождения и других признаков. На европейском континенте общее недопущение дискриминации выражено в Конвенции Совета Европы о защите прав человека и основных свобод (1950), ст. 14 которой устанавливает, что



пользование правами и свободами, признанными в Конвенции, должно быть обеспечено без какой бы то ни было дискриминации по признаку пола, расы, цвета кожи, языка, религии, политических или иных убеждений, национального или социального происхождения, принадлежности к национальным меньшинствам, имущественного положения, рождения или по любым иным признакам. Заметим, что применение процитированной статьи создало богатую практику Европейского суда по правам человека (далее — ЕСПЧ), с разными нюансами — хотя сама ст. 14 не гарантирует самостоятельного права не подвергаться дискриминации, Суд толкует ее таким образом, что даже в случае отсутствия нарушения «материальной» статьи может быть установлено нарушение ст. 14 [Бартенев Д.Г., 2019:43-66].

В Европейском Союзе дискриминация запрещена Директивой Совета 2000/43/ЕС от 29 июня 2000 года, закрепившей осуществление принципа равного обращения с лицами независимо от их расовой принадлежности или этнического происхождения и, в конкретном контексте занятости, — Директивой Совета 2000/78/ЕС от 27 ноября 2000 года, устанавливающей общие рамки равного обращения при трудоустройстве. Определение дискриминации, содержащееся в обеих директивах 2000 г., носит широкий характер, поскольку охватывает не только понятия прямой и косвенной дискриминации, но также включает в себя концепцию преследования. В более прикладном плане для Суда ЕС дискриминация может состоять в применении разных норм к сравнимым ситуациям или одной и той же нормы к разным ситуациям¹.

В некоторых национальных юрисдикциях закреплена общая антидискриминационная система защиты на конституционном уровне. Например, ст. 3 IV Конституции Бразилии называет одной из основных целей Республики содействие благосостоянию всех людей без ущемлений по причине происхождения, расы, пола, цвета кожи, возраста и любых других форм дискриминации, а ст. 7 запрещает любую дискриминацию в отношении заработной платы и критериев найма работников-инвалидов. Кстати, в Конституции Российской Федерации также запрещена дискриминация относительно вознаграждения за труд (ст. 37), равно как дискриминация в

1. CJCE 17 juin 1998, Kathleen Hill, aff. C-243/95, rec. 1998, p. 1-3739



отношении социальной интеграции инвалидов (ст. 114). Общий запрет дискриминации сформулирован в ст. 19 Конституции РФ.

Наш нормативный экскурс имел целью не полноту, а, скорее, должную иллюстративность проблематики дискриминации. Мы выяснили, что данная тема заняла самостоятельное место как на международном, так и на национальном уровне, а также имеет «секторальные предпочтения». Тем, что право на недискриминацию базируется на равенстве, объясняется четкая прослеживаемость общего принципа равенства при анализе права на недискриминацию (т.е. рассмотрение аналогичных ситуаций одним и тем же способом и рассмотрение различных ситуаций по-разному). При этом значительную долю занимают конкретные положения, разработанные в специальном законодательстве в области борьбы с дискриминацией, касающиеся определенных охраняемых оснований (например, возраст, раса, пол, религия, сексуальная ориентация и т.д.), а также конкретная область применения (рынок труда, профессиональная подготовка, образование, социальное обеспечение, здравоохранение, доступ к товарам и услугам, уголовная ответственность). Таким образом, право на недискриминацию — симбиоз, в котором каждый элемент имеет свое место и значение. При этом право на недискриминацию не абсолютно хотя бы потому, что закон правового государства может санкционировать дискриминацию, если предполагается ее необходимость для «выравнивания» прав и возможностей разных категорий лиц. Что меняют в проанализированном status quo цифровые технологии?

2. Профилирование как дискриминационная практика

2.1. Дискриминационные риски профилирования

Профилирование (профайлинг) является конкретным методом интеллектуального анализа данных. В этой перспективе профилирование рассматривается как автоматизированный либо полуавтоматизированный процесс изучения больших наборов данных в целях построения классов или категорий характеристик [Bosco F. et al., 2014: 4]. В правовой литературе существуют разные подходы к определению профилирования. Так, профилирование определяется как объединение набора характеристик (профиля) об индивидууме или организации и



последующее обращение с этим лицом или организацией сообразно таким характеристикам [Bygrave L. A., 2002: 301]. Иногда в определении акцентируется аспект идентификации: профилирование как «процесс» обнаружения паттернов в различных данных, которые могут использоваться для идентификации лица (индивидуальный или групповой) и/или применение профилей (наборов коррелирующих данных) для индивидуализации субъекта или для его идентификации как члена группы [Hildebrandt M., 2009:275].

И в том, и в другом подходах выделены важные компоненты профилирования — поиск информации по определенным критериям, связь этой информации с конкретным лицом, использование ее для идентификации лица или для обращения с лицом неким заданным образом. В плане работы с информацией профилирование представляет собой отход от идеи, что знание является результатом проверенной гипотезы. По сути, профилирование само генерирует гипотезы: корреляции как таковые становятся релевантной информацией. В силу изложенного профилирование называют «гламурной технологией»: оно зиждется на идее о том, что люди могут получить непредвиденные знания, которые позволяют принимать лучшие решения [Bosco F. et al., 2014: 7, 12].

Наше исследование концентрируется на том, как в процессе профилирования способна проявить себя дискриминация. Дело в том, что при профилировании обрабатываются как персональные, так и неперсональные данные. Из этого посыла сразу ясно, что основной конфликт профилирования в плоскости законодательства отмечается с законодательством о защите данных. Но это конкретный аспект, который в целом нормативно уже регулируется, пусть и неидеально. Гораздо более проблемной становится обработка данных в условиях незнания граждан о том, что их данные обрабатываются и используются для принятия решений. Таким образом, в более общем плане профилирование «сталкивается» с либеральной демократией, прежде всего, потому, что провоцирует информационную асимметрию. В общем-то, асимметрия знаний — явление распространенное, но именно в технологиях профилирования она достигает нового пика, именно потому, что в большинстве случаев граждане не знают о собираемой и распространяемой информации и о том, как ее можно использовать в будущем.

Собственно, именно эти риски и были названы основными в социологическом опросе, проведенном в 2014 году среди



профессионалов — национальных органов по защите персональных данных европейских стран. По их обобщенному мнению, в связи с профилированием возникают проблемы для основных прав и свобод граждан в целом (частная жизнь и защита данных, самоопределение, достоинство, личная неприкосновенность, свобода слова и передвижения), а также проблемы в связи с отсутствием транспарентности и осведомленности о существовании и проведении профилирования [Bosco F. et al., 2014: 24].

Очевидно, что такие технологии, как профилирование, не оставляют большого пространства для автономии и самоопределения личности. Изначально заложенное противоречие между закрытостью частной жизни и открытостью публичной сферы обретает новые грани — частная жизнь тяготеет к непрозрачности, а условием защиты данных является открытость. Вопрос — как в таких условиях защитить частную жизнь? — становится трудноразрешимым.

Данный поворот заставляет заново взглянуть на теорию информационного самоопределения [Rouvroy A., Poullet Y., 2009: 51] и, возможно, переоценить ее. В случае с профилированием она приобретает совершенно конкретное звучание. Теория информационного самоопределения базируется на постулате, что индивидуум должен осуществлять контроль за данными и информацией, производимыми им. Возможно, всесторонняя поддержка и внедрение принципов информационного самоопределения в политику данных позволит разделить ответственность обработчиков и обладателей данных, сделает их отношения более горизонтальными.

Прежде чем перейти собственно к дискриминации при профилировании, приведем конкретный пример. Вид профилирования, получивший название «редлайнинг» (redlining), уже напрямую запрещен законодательством США. Редлайнинг используется для выявления практики отказа от товаров и услуг в конкретных районах, отмеченных красной линией на карте. Вследствие расовой сегрегации или увеличения демографической концентрации люди, аналогичные по социальному классу, условиям занятости и даже национальности или проживающие в конкретном районе, зачастую могут принадлежать к определенной расовой группе или этническому меньшинству [Bosco F. et al., 2014: 20]. Возможная в таких случаях дискриминация и послужила основанием для запрета редлайнинга.



2.2. Генетическое профилирование в контексте дискриминации

Пожалуй, самой яркой моделью дискриминационного профилирования служит генетическое профилирование. Генетический профиль способен раскрыть информацию не только о человеке, но и о его родственниках, включая их текущий и будущий профиль здоровья и ранее неизвестные семейные отношения [Feldman E.A., Quick B., 2020: 334]. Разумеется, диапазон использования подобного рода информации чрезвычайно широк, что неминуемо создает риски, в том числе дискриминационные.

Поскольку в ряде стран уже сконструировано специальное законодательство в рассматриваемой сфере, полезно обратиться к конкретным иллюстрациям. В законодательстве США (Genetic Information Nondiscrimination Act (GINA) (2008) преследуются две явные цели — предотвращение генетической дискриминации в медицинском страховании и предотвращение генетической дискриминации на рабочем месте. Медицинским страховщикам запрещено требовать прохождения генетического тестирования для оформления страховки или принимать решения о возмещении вреда на основании генетической информации. Аналогично работодателям запрещено требовать прохождения генетического тестирования и принимать решения о найме, увольнении и пр. на основе генетической информации. Заметим, что речь идет именно о медицинском страховании, а не о страховании жизни. К тому же законодательство о генетической дискриминации не распространяется на военных [Feldman E.A., Quick E., 2020: 336].

В Эстонии существует общий Закон о равном обращении — Equal Treatment Act (2017)², который действует как для государственного, так и для частного сектора. Им запрещается дискриминация по признаку национальности (этническое происхождение), расы, цвета кожи, религии или других убеждений, возраста, инвалидности или сексуальной ориентации. Что касается конкретно генетической дискриминации, то эстонский Закон об исследованиях генов человека (Estonian Human Genes Research Act) (2014)³ устанавли

2. Estonian Equal Treatment Act—RT I, 26 April 2017, 9. Available at: URL: <https://www.riigiteataja.ee/en/eli/503052017002/consolide> (дата обращения: 17.02.2022)

3. Estonian Human Genes Research Act—RT I, 14 March 2014, 30. Available at: URL: <https://www.riigiteataja.ee/en/eli/518062014005/consolide> (дата обращения: 17.02.2022)



вает общий запрет на генетическую дискриминацию, сопровождаемый двумя конкретными запретами в сфере страхования и занятости. Более конкретно, работодателям запрещено собирать генетические данные о сотрудниках или требовать от сотрудников и соискателей должности предоставлять образцы тканей или ДНК. Работодателям запрещено навязывать дискриминационные условия труда и оплаты труда для людей с различными генетическими рисками [Pormeister K., 2020:179]. В отличие от США, эстонский запрет на генетическое тестирование в страховании не ограничивается только медицинским страхованием: всем страховым компаниям запрещается требовать генетические данные и обуславливать условия страхования генетическими факторами.

Конечно, в идеале необходим такой правовой режим генетической информации, который позволял бы извлекать из нее выгоды и одновременно не подвергать рискам ее владельцев. Как утверждают исследователи, в Канаде особый правовой режим, установленный Законом о генетической недискриминации (2017), позволяет лицам проходить генетическое тестирование и пользоваться его преимуществами, не подвергаясь при этом рискам, связанным со страхованием, или дискриминации в области занятости по признаку генетического статуса. При этом указывается, что права и обязанности страховщиков и работодателей по отношению к членам семьи или родственникам нуждаются в уточнении, в частности, в вопросах добровольного раскрытия информации [Khoury L. et al, 2020:47].

Однако, как это часто бывает, за, казалось бы, четкими запретами скрываются разнообразные условия. Представляется возможным выделить три проблемы.

Во-первых, с учетом практики легальная дискриминация иногда может быть полезной и даже необходимой. Например, как утверждает эстонский исследователь Nomper [Nomper A., 2001: 120], работодатель должен иметь право требовать прохождения генетического тестирования как до поступления на работу, так и во время трудовых отношений в той мере, в какой такое тестирование необходимо для определения существования генетического заболевания, которое может привести к выполнению должностных обязанностей с риском для здоровья работника или других сотрудников и иных лиц, активов работодателя или клиентов. Так, проявление болезни Хантингтона (симптомами которой являются непроизвольные



движения) создает потенциальный риск для пилота самолета и его пассажиров.

Во-вторых, в литературе обращается внимание на сложности защиты прав в контексте генетической дискриминации. В частности, по эстонскому законодательству одним из заметных процедурных преимуществ общего Закона о равном обращении (который не распространяется на генетическую дискриминацию) является распределение бремени доказывания. Общий закон устанавливает совместное бремя доказывания, оговаривая, что заявление лица в суд, Комитет по трудовым спорам или Уполномоченному по вопросам гендерного равенства и равного обращения основывается на фактах, из которых можно предположить, что дискриминация произошла. Затем бремя доказывания перекладывается на ответчика, который в ходе разбирательства должен сам доказать, что нарушения принципа равного обращения не было. Если ответчик отказывается представить такие доказательства, это приравнивается к признанию дискриминации [Pormeister K., 2020: 189]. По Закону об исследованиях генов человека доказывать дискриминацию должен заявитель (как и то, что работодатель использовал генетическую информацию), что, конечно, крайне сложно на практике.

В-третьих, нельзя отрицать предиктивного потенциала генетической информации. Пока генетическая информация не является точной на все 100%. Однако если она станет таковой (в том смысле, что станет возможным со 100%-ной уверенностью установить индивидуальный риск), мы столкнемся с трудным выбором между отказом от использования предсказательной силы генетической информации и явной дискриминацией людей на основе качеств, над которыми они не имеют никакой власти и контроля [Feldman E.A., Quick E., 2020: 351]. Предиктивный потенциал данных, не только генетических, заслуживает более пристального рассмотрения.

3. Предиктивность и персонализация: позитивные и негативные стороны

Одним из преимуществ алгоритмов, используемых искусственным интеллектом, называют предиктивность, возможность предсказать и предугадать дальнейшее развитие событий. Это звучит довольно безобидно, если речь идет о построении планов развития экономики на государственном



уровне, о маркетинговых прогнозировании частных компаний. Но совсем другое дело, когда предиктивные способности искусственного интеллекта используются для юридической оценки поведения конкретного человека. Как утверждают исследователи, именно в целях расследования преступлений сексуального характера еще в середине XX в. стало использоваться профилирование. Наибольшую известность профайлинг получил в США и странах Западной Европы в рамках расследований серийных убийств. В нашей стране подобная деятельность получила свое название — криминалистическое портретирование [Романовский В.Г., 2020:47].

В качестве иллюстрации обратимся к широко используемой американскими судами практике алгоритмического прогнозирования поведения человека, главным образом в рамках уголовного правосудия. Уже многие американские штаты разрешили использовать алгоритмы для принятия решения о предварительном заключении или условно-досрочном освобождении. В частности, речь идет о приложениях PSA (Public Safety Assessment) или COMPAS (Courcional Offender Management Profiling for Alternative Sanctions). Критика подобного алгоритмического прогнозирования поведения человека основана на том, что предиктивный результат вовсе не дает истины, верной на все 100%: его средняя предиктивность оценивается лишь в 65%. Соответственно, тот факт, что судья при вынесении решения опирается на прогнозную оценку алгоритма, представляет собой проблему с точки зрения права на справедливое судебное разбирательство, в том числе в связи с тем, что судья может пытаться выполнить рекомендацию алгоритма почти дословно и систематически [G'sell F., 2020: 302].

Естественно, что судебные решения, принятые подобным образом, неоднократно становились объектом обжалования. Так, в деле Loomis, которое было рассмотрено в 2017 году Верховным судом штата Висконсин⁴, поднимался вопрос о допустимости применения алгоритмической предикции. Апелляция гражданина была отклонена Верховным судом штата Висконсин, который не посчитал, что ранжирование отдельных лиц по группам профиля является серьезной проблемой в свете принципа индивидуализации наказания. Суд счел, что права гражданина Loomis соблюдены, если проведенная алгоритмом оценка не является единственной основой для принятого

4. State of Wisconsin c/ Loomis, 881 N.W.2d 749 (Wis. 2016).



судебного решения. Судьи могут даже не принимать во внимание алгоритмическую оценку и отдавать предпочтение иным данным. Верховный суд штата Висконсин также уточнил, что судьи, прибегающие к такому виду инструмента, должны всегда четко упоминать в своих решениях другие факторы, лежащие в основе их решения.

Конечно, такое объяснение частично снимает остроту проблемы, если только не учитывать склонность людей довериться машине. Человеческий мозг и вправду кажется не таким мощным, быстрым, вместительным, как искусственный интеллект. С той лишь разницей, что только человеческий мозг способен понять мотивы совершенного преступления, найти им не формальное, а контекстное объяснение, учесть конкретные обстоятельства, зачастую очень личного, интимного характера. Для машины же все эти обстоятельства — просто данные в виде нулей и единиц.

Кроме того, пресловутая проблема «черного ящика» не позволяет установить с достоверностью, почему система приняла именно такое, а не иное решение. Опять-таки частичное решение проблемы — сделать алгоритм полностью транспарентным и доступным, но даже в этом случае понять его функционирование вряд ли возможно полностью. Кроме того, если говорить конкретно о системах, используемых в США, то методология алгоритма, включая значение каждой переменной в вычислении, защищается в качестве тайны дела и не раскрывается ни судье, ни осужденному. Разумеется, подобная непрозрачность создает массу поводов для беспокойства.

В Европе традиционно с большой внимательностью подходят к вопросам прав человека и открытости власти. Например, во Франции использование алгоритмов для оценки потенциальной опасности лица будет не совсем законно. Статья 47 Закона об информатике, картотеках и свободах (1978)⁵ устанавливает, что судебные решения, предусматривающие оценку поведения человека, не могут основываться на автоматизированной обработке персональных данных, имеющей целью оценку отдельных аспектов личности. Правда, некоторая подвижка произошла с принятием Закона № 2019- 222 от 23 марта 2019 г., которым введена новая статья в Закон № 2016-1547 от 18 ноября 2016 г. о модернизации правосудия XXI века. Она предусматривает, что процедуры примирения, посредничества и

5. Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Available at: URL: <https://www.legifrance.gouv.fr/loda/id/> (дата обращения: 17.02.2022)



третейского разбирательства, проводимые онлайн, не могут основываться исключительно на алгоритмической или автоматизированной обработке персональных данных. Таким образом, использовать алгоритмы для оценки поведения человека в рамках судебного процесса не запрещено полностью, но результаты анализа на основе автоматизированной обработки не могут служить единственной основой для принятия решения.

Отдельный вопрос — на базе каких ресурсов развивать алгоритмическую обработку данных. Главным источником сведений для обучения и работы искусственного интеллекта служат ранее вынесенные судебные решения. Во Франции решили массово распространять судебные решения с подачи Закона № 2016-1321 от 7 октября 2016 года о цифровой республике, который предусматривает бесплатное предоставление всех данных, полученных от государственных органов, и, следовательно, решений всех судов. Закон № 2019-222 от 23 марта 2019 года о программе правосудия несколько изменил действовавший порядок, предусмотрев, что решения, вынесенные французскими судами, будут предоставлены бесплатно для общественности в электронном виде. Как только эта реформа будет полностью осуществлена, должны появиться обширные базы данных судебных органов, создающие основу работы с алгоритмами машинного обучения. Тогда алгоритм сможет предугадать, каким может быть решение суда с учетом всех решений, принятых в прошлом. Однако алгоритм не сможет предложить правовых аргументов, обосновывающих это предположение, поскольку в системе континентального права (где общее правило применяется к конкретному случаю) учет прошлых решений не является достаточным для рассмотрения дела за исключением особо простых и повторяющихся споров. Единственное, что могут сделать алгоритмы, — это обрабатывать повторяющиеся ситуации, такие, например, как нарушение правил парковки или урегулирование конфликтов на платформе электронной коммерции. Алгоритмы не в состоянии рассматривать дела, имеющие определенную степень сложности.

Еще одна французская особенность — Закон 2019 года запретил именное профилирование судей (это значит, что в алгоритме не должны использоваться идентификационные данные судьи и судебного секретаря). При этом можно анализировать не только судебную практику по всей Франции, но и практику конкретного суда, сравнивая судебные решения в целом. Тем не менее, несколько французских стартапов (таких,



как Case Law Analytics и Predictice) предлагают инструменты предсказания, позволяющие предвидеть возможные результаты судебной процедуры. Эти программные инструменты, в основном, обеспечивают оценку размера денежных средств, которые могут быть присуждены судом (ущерб и проценты, алименты, компенсационные пособия). Поэтому некоторые специалисты предлагают использовать термин «количественное правосудие» (justice quantitative), а не «предсказывающее правосудие» (justice predictive), поскольку математические инструменты используются для моделирования решений в области правосудия, создания инструментов анализа и прогнозирования и количественной оценки судебных случаев [G'sell F., 2020: 307].

Таким образом, развитие приложений по решению рутинных правовых проблем, которые будут давать результаты на основании прошлых судебных решений, особенно актуально в странах общего права, где юристы систематически ссылаются на предыдущие решения. Напомним, что Россия принадлежит к континентальной системе права, и судебная практика даже по одному виду споров зачастую оказывается противоречивой и не совпадающей в разных регионах страны.

4. Дискриминационная обработка данных искусственным интеллектом

Существующие исследования дискриминационных рисков при обработке данных так или иначе выделяют несколько категорий недостатков. В считающейся основополагающей статье Barocas и Selbst [Barocas S., Selbst A.D., 2016: 677-693] выделяют пять способов, с помощью которых решение, принимаемое искусственным интеллектом, может непреднамеренно привести к дискриминации. В широком плане проблемы связаны с: определением целевой переменной и классовых меток; маркировкой тренировочных данных; сбором тренировочных данных; выбором характеристик; выбором проксиданных. Кроме того, шестым способом дискриминации является преднамеренное использование систем искусственного интеллекта в дискриминационных целях.

Рассмотрим виды недостатков, которые способны привести к дискриминационной обработке данных, а именно: неправильный сбор данных, агрегирование ошибочных данных, и невосприимчивость к нормативным установкам.



1. Сбор данных с дефектами. Общая причина недостаточных или даже дискриминационных результатов в решениях искусственного интеллекта — создание ошибочных обучающих данных. Смещенные входные данные с большей вероятностью приведут к смещенным выходным данным. Такое смещение на входе может быть вызвано либо недостаточной, либо избыточной представленностью данных.

Приведем довольно типичный пример. В паспортных органах Новой Зеландии используется автоматизированная система проверки фотографий, сделанных для получения паспорта. Система регулярно отклоняет фотографии людей азиатского происхождения, квалифицируя их изображения как не отвечающие требованиям, потому что на них изображены заявители с закрытыми глазами [Regan J., 2016]. Такое предвзятое автоматизированное решение стало прямым результатом ошибочного сбора исходных данных, т.е. система была создана с данными для обучения, собранными в основном от лиц неазиатского происхождения. Иными словами, результаты, выдаваемые искусственным интеллектом, не могут быть лучше, чем данные, которыми он «питался».

Другой пример — приложение Street Bump для смартфона использует геолокацию для мониторинга состояния дорог в городе. Если у меньшего числа бедных людей будет меньше смартфонов, чем у более богатых, то эта категория данных будет представлена недостаточно. Что может привести к тому, что об изношенных дорогах в бедных районах будет реже сообщаться в наборах данных, а, значит, реже будет производиться их ремонт [Zuiderveen Borgesius F., 2018: 10].

2. Агрегирование ошибочных данных. Еще одной причиной дискриминационного подхода со стороны искусственного интеллекта является неправильное агрегирование данных. В этом случае исходные данные для обучения как таковые репрезентативны, но проявления предвзятости возможны позднее, в процессе обработки.

Во-первых, искусственный интеллект может быть обучен на основе изначальной ручной агрегации данных и воспроизводить предвзятость конкретных лиц, ранее принимавших решения на основе этих данных, т.е. решения, принятые когда-то людьми, становятся для искусственного интеллекта примером. Например, разные дефекты маркировки данных, когда технически невозможно отметить в качестве кредитоспособного человека из этнических меньшинств, который допустил три просрочки



выплат по кредитам, в то время как для «белых» допускаются четыре задержки выплат по кредиту.

Данные для обучения становятся предвзятыми, если они отражают дискриминационные человеческие решения. Подобный пример описан в литературе. Для отбора в один из медицинских колледжей Великобритании, систематически получавший больше заявок на поступление, чем имелось мест, было разработано программное обеспечение, помогающее сортировать заявки. Алгоритм был обучен на основе записей о приеме в колледж в предыдущие годы, когда абитуриентов отбирали в «ручном режиме», т.е. по сути была воспроизведена предшествующая система отбора. В итоге было установлено, что компьютер ставит в невыгодное положение женщин и мигрантов, т.е. в предшествующие годы лица, ответственные за отбор студентов, предвзято относились к этим категориям лиц. Программа не внесла никакой новой предвзятости, а лишь повторила предвзятость, которая уже существовала в системе.

Во-вторых, дефекты могут быть спровоцированы и после подготовительного периода. Например, если выполнить поиск через Гугл женских причесок в деловом стиле и причесок неделового стиля, то поисковик показывает в первом случае в основном фотографии белых женщин со светлыми волосами, и фотографии темнокожих женщин — во втором. Гипотетически, если работодатель разрешает искусственному интеллекту отсканировать заявления на работу и разобраться с ними по критерию «непрофессиональный вид», это способно привести к юридически значимой расовой дискриминации [Tischbirek A., 2020: 106].

3. Невосприимчивость к нормативным установкам. Здесь речь идет о другой, более фундаментальной проблеме принятия решений на основе применения искусственного интеллекта, а именно, о статистике, которая может оценивать будущее только с учетом прошлого. Нормативные же соображения обусловлены оценкой фактов. Это значит, что перевести такие соображения на язык, понятный искусственному интеллекту, может быть весьма проблематично. В качестве примера в литературе приводится одно из самых значимых антидискриминационных решений Суда ЕС за последнее десятилетие — по делу Бельгийской ассоциации потребителей Test-Achats. В этом деле Суд заявил, что страховые тарифы, различающиеся в зависимости от пола клиента, нарушают принцип недискриминации ЕС⁶. Постановление Суда ЕС — пример того,



как нормативное по своей природе предписание должно бороться с фактами современности на благо иной, будущей реальности.

В частности, были исследованы разнообразные статистические данные, говорящие о том, что мужчины гораздо чаще попадают в автомобильные аварии. И опять же, средняя продолжительность жизни женщин выше. Наконец, женщины обычно сталкиваются с большими расходами, связанными с беременностью. Следовательно, можно ожидать, что страхование автомобиля и страхование жизни должны обходиться дороже для мужчин, чем для женщин, в то время как логично ожидать, что тарифы медицинского страхования будут выше для женщин, нежели для мужчин.

Иными словами, если цифры правильные, то статистические знания призывают страховые компании использовать пол в качестве актуарного фактора. Но в правовом государстве указанные статистические знания и их последствия подчиняются политической оценке и, в конечном счете, юридическому структурированию. Принимаются соответствующие нормы — так, можно распределить расходы на беременность поровну между мужчинами и женщинами для обеспечения гендерных тарифов медицинского страхования; защита материнства — это даже конституционное обязательство во многих странах [Tischbirek A., 2020: 107].

Однако нормативные выводы, такие как требование об унисекс-тарифах, не столь устойчивы, когда в игру вступает искусственный интеллект. Так, простое исключение чувствительного параметра (в данном случае пола) не является перспективным, поскольку искусственный интеллект может «узнать» о поле кандидата, соотнести его со статистикой и предложить соответствующий тариф страхования. Постоянно меняющаяся, динамическая конфигурация алгоритма делает чрезвычайно трудным — или даже невозможным — ретроспективное определение всех решающих факторов при расчете конкретного страхового тарифа [Tischbirek A., 2020:108-109].

6. CJEU C-236/09 'Association belge des Consommateurs Test-Achats ASBL et al. v. Conseil des ministres' (1 March 2011), 2011 E.C.R. 773, paras 30–32.



5. Поиск правовых решений

Риски, в том числе дискриминационные, к которым может вести применение искусственного интеллекта в принятии решений, признаются во всем мире. В связи с этим интересно сопоставить различные подходы к урегулированию на современном этапе практики использования искусственного интеллекта с точки зрения прав человека.

Наиболее щепетильно к правам человека относится Европа, которая довольно последовательно движется по пути жесткой регламентации использования искусственного интеллекта. В Белой книге по искусственному интеллекту (а European approach to excellence and trust⁷) Европейский Союз в феврале 2020 уделил значительное внимание этике применения искусственного интеллекта, делая акцент на человекоцентричности и риск-ориентированном подходе к использованию технологий. В апреле 2021 г. появился проект Европейского регламента по искусственному интеллекту⁸, в котором соблюдение и эффективное применение существующего законодательства об основных правах имеют ключевое значение. Особую озабоченность вызвала возможность использования технологии не по назначению, с возникновением новых и мощных инструментов для манипулирования, эксплуатации и социального контроля. Поскольку подобная практика противоречит ценностям Союза и подрывает право на недискриминацию, право на защиту частной жизни и персональных данных, ее предполагается запретить. Так, статьей 5 проекта регламента запрещены следующие виды практики в области искусственного интеллекта:

маркетинг, ввод в эксплуатацию или применение системы искусственного интеллекта, использующей сублиминальные методы воздействия на подсознание человека, которые причиняют или могут причинить физический или психологический вред этому человеку или третьему лицу;

маркетинг, ввод в эксплуатацию или применение системы искусственного интеллекта, которая использует любые уязвимости, связанные с возрастом или физическими или умственными недостатками определенной группы лиц, для

7. Available at: URL: <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:52020DC0065> (дата обращения: 17.02.2022)

8. Available at : URL : <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206> (дата обращения: 17.02.2022)



существенного изменения поведения члена этой группы таким способом, который причиняет или может причинить физический или психологический вред этому или третьему лицу;

маркетинг, заказ или использование государственными органами или от их имени систем искусственного интеллекта, предназначенных для оценки или ранжирования надежности отдельных лиц в течение определенного периода времени на основе их социального поведения или известных либо прогнозируемых личностных характеристик. Социальный балл (оценка) при этом влечет предвзятое или неблагоприятное отношение к определенным лицам, группам лиц в тех аспектах, которые не связаны с контекстом получения данных, или же такое предвзятое отношение является неоправданным или несоразмерным;

использование систем дистанционной биометрической идентификации в режиме реального времени в общественных местах в целях правоохранительной деятельности (за исключением поименованных в проекте регламента случаев, таких как целенаправленный поиск конкретных потенциальных жертв преступлений, включая пропавших детей; предотвращение конкретной, существенной и неминуемой угрозы жизни или физической безопасности людей или предотвращение террористического акта; обнаружение, определение местонахождения, идентификация или преследование исполнителя или подозреваемого в совершении уголовного преступления).

Помимо этого проект регламента градирует системы искусственного интеллекта по степеням риска. В статье 6 определены правила отнесения систем искусственного интеллекта к категории высокого риска (это, например, скоринг при найме на работу, онлайн-биометрия, предиктивные системы в судопроизводстве). Таким образом, разновидность использующейся в Китае системы социального скоринга (оценка и прогнозирование социального поведения лица) скоро будет полностью запрещена в Европе, а использование систем распознавания лиц — строго ограничено.

На американском континенте, несмотря на схожую озабоченность проблемами предвзятости и потенциальной дискриминации со стороны искусственного интеллекта, о прямых запретах речь пока не идет. Специалистами Национального института стандартов и технологий (National Institute of Standards and Technology, NIST) в июне 2021 года



подготовлен проект предложений, изложенных в докладе «A Proposal for Identifying and Managing Bias in Artificial Intelligence»⁹. Отталкиваясь от определения Международной организации стандартизации (ISO), в котором под статистическим смещением, предвзятостью понимается степень, в которой эталонное значение отклоняется от истины, в докладе подчеркивается, что отклонение от истины может быть как положительным, так и отрицательным, оно может способствовать вредным или дискриминационным результатам, но может быть полезным. С точки зрения общества, предвзятость часто связана с ценностями и рассматривается через двойную призму дифференцированного обращения или диспропорционального воздействия — ключевых юридических терминов, относящихся к прямой и косвенной дискриминации соответственно.

Предполагается, что предвзятость присутствует во всех системах искусственного интеллекта, а собственно проблема заключается в выявлении предвзятости, ее измерении и управлении ею. Существующие отраслевые подходы в этом мало помогают — принято классифицировать предвзятость по типу (статистическая, когнитивная) или по сфере использования (прием на работу, здравоохранение и т.д.) — и не могут обеспечить широкой перспективы, необходимой для эффективного управления предвзятостью как явлением, зависящим от контекста. Соответственно, в докладе предпринята попытка преодолеть эту узость, и предлагается подход к управлению и снижению воздействия вредных предубеждений в различных контекстах. Для этого, отмечают разработчики, нужно использовать ключевые моменты на этапах жизненного цикла искусственного интеллекта для оптимального выявления и управления предубеждениями.

Основные выводы в докладе сводятся к следующему: предвзятость не является ни новым, ни уникальным явлением для искусственного интеллекта; цель заключается не в нулевом риске, а в выявлении, понимании, измерении, управлении и уменьшении предвзятости; необходимы стандарты и руководства, определяющие терминологию, способы измерения и оценки предвзятости; необходимы методы снижения предвзятости, которые являются гибкими и могут применяться в различных контекстах, независимо от отрасли; NIST планирует

9. Available at : URL : <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270-draft.pdf> (дата обращения: 17.02.2022)



разработать требования к надежному и ответственному использованию искусственного интеллекта при участии широкого круга заинтересованных сторон, чтобы гарантировать, что стандарты и практика отражают точки зрения, которые традиционно не учитываются при разработке технологий; NIST разработает дополнительное руководство по обеспечению, управлению и совершенствованию практики, а также методы улучшения коммуникации между различными заинтересованными сторонами.

Таким образом, американский вариант представляет собой развитие преимущественно технического регулирования без жесткого законодательства. Примерно в том же ключе выстраивается понимание проблематики в Австралии. Комиссия по правам человека Австралии (the Australian Human Rights Commission) в марте 2021 года представила доклад «Human Rights and Technology»¹⁰. Принципы этики искусственного интеллекта, разработанные правительством Австралии, гласят, что системы искусственного интеллекта не должны приводить к несправедливой дискриминации. В принципах признается ответственность тех, кто участвует в принятии решений на основе искусственного интеллекта, чтобы гарантировать, что эти решения соответствуют антидискриминационным законам.

В качестве решения проблемы алгоритмической предвзятости в докладе предлагается расширить подотчетность принятия решений на основе искусственного интеллекта. Кроме того, может возникнуть необходимость в более целенаправленной реформе антидискриминационного законодательства, чтобы решить такие вопросы, как воздействие зависимости от конкретного фактора или комбинации факторов на защищенные признаки, и каким образом это может быть связано с правовыми концепциями прямой и косвенной дискриминации. Также необходимо изучить, как конкретные методы деятельности искусственного интеллекта, такие как предиктивные алгоритмы машинного обучения, могут привести к новым формам дискриминации посредством установления связей на основе множества точек данных. В итоге констатируется необходимость разработки руководства для правительственных и неправительственных организаций по соблюдению федеральных антидискриминационных законов при принятии решений на основе искусственного интеллекта.

10. Available at: URL: https://tech.humanrights.gov.au/sites/default/files/2021-05/AHRC_RightsTech_2021_Final_Report.pdf (дата обращения: 17.02.2022)



Как видим, в силу инновационного характера искусственного интеллекта пока ни одно государство не выработало окончательной позиции и четких правил его антидискриминационного функционирования. Напомним, что в России использование искусственного интеллекта не регулируется специально, существуют только нормы законодательства о защите персональных данных, которые необходимо соблюдать, в том числе при использовании цифровых технологий. При этом в российской науке проблемы дискриминации при применении искусственного интеллекта практически не поднимаются.

Наконец, для полноты картины обратимся к международному уровню. Совет по правам человека ООН (Human Rights Council) принял по итогам своей 47-й сессии Резолюцию о поощрении, защите и осуществлении прав человека в Интернете¹¹ (июль 2021 г.). Данный документ ООН связан с конкретной ситуацией блокировок Интернета в Мьянме и разработан с целью осуждения практики блокировок Интернета и онлайн-цензуры; при этом в тексте многократно упоминается дискриминация. Резолюция призывает все государства активизировать усилия по преодолению цифрового разрыва, включая гендерный цифровой разрыв, и расширить использование информационно-коммуникационных технологий в целях содействия осуществлению прав человека для всех, в том числе путем содействия созданию благоприятной онлайн-среды, безопасной и способствующей участию всех без какой-либо дискриминации и с учетом интересов лиц, сталкивающихся с систематическим неравенством (systemic inequalities). Системный характер дискриминации, действительно, является наиболее опасным последствием применения искусственного интеллекта, в силу чего актуализируется необходимость полноценного урегулирования и контроля качества данных, используемых для обучения искусственного интеллекта.

Заключение

Обработка данных искусственным интеллектом, с одной стороны, повторяет, имитирует человеческую деятельность (и в этом смысле воспроизводит все человеческие предвзятости), но, с другой стороны, способна игнорировать нормативные препоны

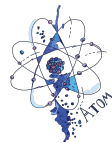
11. Available at: URL: https://d-russia.ru/wp-content/uploads/2021/07/un_hrc_internet_english.pdf (дата обращения: 17.02.2022)



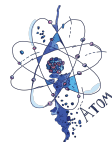
в виде декларативных призывов к недискриминации. Это означает, что законодательство, которое в будущем сложится в отношении использования искусственного интеллекта, должно среди прочего учитывать и нейтрализовать риски дискриминации, сочетая при этом правовые и технические подходы. Предлагаемые параметры и критерии алгоритмической обработки должны пройти правовую экспертизу на предмет возможной дискриминации, это относится как к отбору данных для обучения, так и непосредственно к разработке алгоритма. Вероятно, потребуется специальная сертификация со стороны государства разрабатываемых алгоритмов (она может быть шире по предмету, но антидискриминационная оценка должна стать составной частью). Наконец, самое сложное — приоткрыть «черный ящик», сделать алгоритмы понятными и доступными, в том числе для судебного обжалования принятых на их основе конкретных решений.

Список источников

1. Бартенев Д.Г. Запрет дискриминации: практический обзор подходов Европейского суда по правам человека // Международное правосудие. 2019. № 1. С. 43-66.
2. Лапаева В.В. Правовой принцип формального равенства // Журнал российского права. 2008. № 2. С. 67-80.
3. Нерсесянц В.С. К праву. О происхождении равенства (из неопубликованного) // История государства и права. 2009. № 17. С. 2-7.
4. Романовский В.Г. Профилирование террористов и конституционная защита прав человека // Конституционное и муниципальное право. 2020. № 10. С. 46-50.
5. Barocas S., Selbst A.D. Big Data's disparate impact. *California Law Review*, 2016, vol. 104, pp. 671-732.
6. Iosco E, Creemers N., Ferraris V. et al. Profiling Technologies and Fundamental Rights and Values: Regulatory Challenges and Perspectives from European Data Protection Authorities / S. Gutwirth, R. Leenes, P. de Hert (eds.). *Reforming European Data Protection Law*. Berlin: Springer, 2014, pp. 3-33.
7. Bygrave LA. *Data protection law: Approaching its rationale, logic and limits*. The Hague: Kluwer Law International, 2002, 456 p.



8. Feldman E.A., Quick E. Genetic Discrimination in the United States: What State and National Government Are Doing to Protect Personal Information / Khoury L, Blackett A., Vanhonnaeker L. (eds.). Genetic Testing and the Governance of Risk in the Contemporary Economy. Berlin: Springer, 2020, pp. 331-354.
9. G'sell F. Les prog res a petits pas de la «justice predictive» en France. ERA Forum, 2020, vol. 21, pp. 299-310.
10. Hildebrandt M. Profiling and AML / K. Rannenberg, D. Royer, A. Deuker (eds.). The Future of Identity in the Information Society. Challenges and Opportunities. Heidelberg: Springer, 2009, pp. 273-310.
11. Khoury L, Blackett A., Vanhonnaeker L. Legal Aspects of Genetic Testing Regarding Insurance and Employment. General Report / Khoury L. et al (eds.). Genetic Testing and the Governance of Risk in the Contemporary Economy, lus Comparatum — Global Studies in Comparative Law 34. Berlin: Springer, 2020, pp. 3-67.
12. Mine M. Les concepts de discrimination directe et indirecte. ERA Forum, 2003, vol. 4, pp. 30-44.
13. Geenitestiide oiguslikust regulatsioonist. Juridica II, 2001, pp. 113-123.
14. Pormeister K. The Prohibitions Against Genetic Discrimination in Estonia / Khoury L, Blackett A., Vanhonnaeker L. (eds.). Genetic Testing and the Governance of Risk in the Contemporary Economy. Berlin: Springer, 2020, pp. 179-191.
15. Regan J. New Zealand passport robot tells applicant of Asian descent to open eyes, 2016. Available at: URL: [reuters.com/article/us-newzealand-passport-error/new-zealand-passportrobot-tells-applicant-of-asian-descent-to-open-eyes-idUSKBN-13W0RL](https://www.reuters.com/article/us-newzealand-passport-error/new-zealand-passportrobot-tells-applicant-of-asian-descent-to-open-eyes-idUSKBN-13W0RL) (дата обращения: 17.02.2022).
16. Rouvroy A., Poullet Y. The right to informational self-determination and the value of self-development. Reassessing the importance of privacy for democracy / S. Gutwirth et al. (eds.). Reinventing Data Protection? Dordrecht: Springer, 2009, pp. 45-76.
17. Tischbirek A. Artificial Intelligence and Discrimination: Discriminating Against Discriminatory Systems /Th. Wischmeyer, T. Rademacher. Regulating Artificial Intelligence. Berlin: Springer, 2020, pp. 103-121.



18. Zuiderveen Borges!us F. Discrimination, intelligence artificielle et decisions algorith- miques. Etude a l'intention du Service anti-discrimination du Conseil de l'Europe. Strasbourg: Conseil de l'Europe, 2018, 51 p.