№ 02 (88) 2020 ЦИФРОВАЯ ЭКОНОМИКА

Первые шаги правового регулирования искусственного интеллекта в Европе и России

29.02.2020 СТАТЬИ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

10 октября 2019 г. Президент РФ подписал Указ № 490 «О развитии искусственного интеллекта в РФ», которым одобрил Национальную стратегию развития искусственного интеллекта на период до 2030 года как часть национальной программы «Цифровая экономика РФ». Стратегия установила принципы развития и использования ИИ в России. ***Наталья Ковалева*** *сравнивает принципы, заложенные в российской Стратегии, с европейским правовым регулированием.*

Опыт Европейского Союза

По степени значимости можно выделить четыре направления, в которых в настоящее время развивается правовое регулирование ИИ в Европе. Первое и самое важное направление — деятельность Группы экспертов высокого уровня по ИИ — High-Level Expert Group on AI (AI HLEG) [1], созданной Еврокомиссией в июне 2018 г. Эксперты AI HLEG консультируют Еврокомиссию по вопросам выработки мер для реагирования на средне- и долгосрочные вызовы, готовят рекомендации по политике в сфере ИИ, а также по этическим, правовым и социальным вопросам ИИ, включая его влияние на социально--экономическую сферу. В мандат группы также входят полномочия по координации деятельности Европейского альянса по ИИ, поддержка диалога с многосторонним сообществом, обобщение экспертных мнений и их отражение в аналитических обзорах и докладах для Еврокомиссии. В составе группы — 52 эксперта из стран ЕС с бессрочными полномочиями. По поручению Еврокомиссии AI HLEG в 2019 г. разработала два стратегических документа, которые легли в основу политики ЕС в сфере ИИ.

Руководство по этике для надежного ИИ

В числе первых документов, разработанных AI HLEG, — Руководство по этике для надежного ИИ от 08 апреля 2019 г. Оно определяет, что основа для надежного ИИ — приверженность принципам, базирующимся на основных правах человека: уважение самостоятельности человека; предотвращение вреда; справедливость и объяснимость.

Для создания заслуживающего доверия ИИ необходимы, но не достаточны семь ключевых требований:

- содействие и контроль со стороны человека;
- техническая надежность и безопасность;
- неприкосновенность частной жизни и контроль за использованием персональных данных граждан;
- прозрачность;
- многообразие, недискриминационность и справедливость;
- благополучие общества и сохранение окружающей среды;
- обеспечение отчетности и возможности аудита.

Помимо соблюдения перечисленных требований, необходимо также оценивать их исполнение на протяжении всего жизненного цикла системы ИИ с помощью как технических средств, так и нетехнических методов.

Рекомендации по политике и инвестициям ЕС в сфере ИИ

Второй документ, разработанный Группой экспертов высокого уровня по ИИ, — Рекомендации по политике и инвестициям ЕС в сфере ИИ от 26 июня 2019 г. Документ содержит 33 детализированные основные рекомендации, которые затрагивают регулирование, финансирование и стимулирование развития и применения в ЕС технологий ИИ. Основной упор сделан на социальные и этические вопросы. Эксперты настоятельно рекомендуют правительствам стран не осуществлять с помощью ИИ всеобщий надзор за населением в целях национальной безопасности, не контролировать поведение и финансовое состояние граждан. Рекомендации призывают правительства поощрять

разработку и применение надежных ИИ систем, действующих под контролем человека и законно применяемых. Надежный ИИ в соответствии с рекомендациями экспертов предполагает защиту персональных данных и подконтрольность владельцам этих данных. ИИ системы при взаимодействии с людьми должны идентифицировать себя, чтобы не допустить ситуаций, когда, например, робот при онлайн--общении выдает себя за человека.

Надежной ИИ системой можно назвать беспилотного робота--курьера компании Starship Technologies [2]. Для распознавания дороги он использует видеокамеры, спроектированные таким образом, чтобы не снимать кого бы то ни было выше колена. Тем самым обеспечивается уважение неприкосновенности частной жизни и защита персональных данных.

Второе направление правового регулирования использования технологий ИИ при обработке персональных данных — это деятельность Консультативного комитета «Конвенции 108» Совета Европы. Комитет 25 января 2019 г. принял Руководство по защите персональных данных при использовании ИИ [3].

Руководство по защите персональных данных при использовании ИИ

Руководство помогает создателям политик, разработчикам ИИ, производителям продуктов и поставщикам услуг в обеспечении того, чтобы ИИ приложения не подрывали право на защиту персональных данных. Документ включает в себя следующие положения:

> ИИ-приложения во всех случаях должны в полной мере соблюдать права субъектов данных;

> ИИ-приложения должны допускать для субъектов данных возможность осмысленно контролировать обработку данных и воздействие на физических лиц и общество;

- на всех стадиях обработки, включая сбор данных, разработчики и производители ИИ должны внедрить подход «запроектированной защиты прав человека» (human rights by-design) и избегать любой потенциальной предвзятости (в том числе неумышленной или неявной), риска дискриминации и иных негативных последствий для прав человека и фундаментальных свобод субъектов данных;
- все продукты и услуги должны проектироваться таким образом, чтобы обеспечивалось право физических лиц на то, чтобы существенно затрагивающие их решения не основывались исключительно на автоматизированной обработке данных, без учета их мнения;
- следует информировать субъектов данных о том, что они взаимодействуют с ИИ-приложением. Они должны иметь право получать сведения о логике, лежащей в основе операций обработки данных с применением ИИ, которые затрагивают их интересы;
- должно быть обеспечено право возражать в связи с обработкой на основе технологий, которые воздействуют на мнение и личное развитие физических лиц.

Третье правоформирующее направление — деятельность Рабочей группы по этике и защите данных в ИИ, которая была учреждена 23 октября 2018 г. на 40-й Международной конференции уполномоченных по защите данных и конфиденциальности. Там же была принята Декларация об этике и защите данных в системах искусственного интеллекта [4].

Декларация об этике и защите данных в системах ИИ

Декларация установила принципы защиты персональных данных в ИИ:

- справедливость — технологии ИИ должны быть спроектированы так, чтобы учитывать их влияние на человека и общество в целом, чтобы не ставить в опасность развитие человечества;

- отчетность — создание и использование ИИ должны подкрепляться отчетностью и контролем на государственном уровне;
- прозрачность / открытость — люди и общество в целом понимают, принимают и доверяют новым технологиям и технологическим решениям. При этом для каждой конкретной аудитории необходим свой уровень открытости и информированности, чтобы каждый человек мог понять, как работает ИИ технология;
- ответственность при создании и развитии ИИ — необходимость оценки и документирования ожидаемого влияния ИИ технологий на человека и общество в целом как в начале процесса создания ИИ, так и на протяжении всего жизненного цикла ИИ;
- недискриминация — уважение общечеловеческих ценностей и прав, выявление предвзятостей и их митигация, разработка специальных правил и принципов, которые помогут минимизировать необъективность в ИИ технологиях.

Наконец, четвертое направление — деятельность неправительственных организаций. Здесь лидирующая роль принадлежит Центру по информационной политике под руководством юридической фирмы Hunton Andrews Kurth, предложившему 25 января 2019 г. дополнения к Декларации об этике и защите данных в системах ИИ [5].

Дополнения к Декларации об этике и защите данных в системах ИИ

Поскольку сегодня сложно с осознанной уверенностью предсказать развитие технологий в будущем, нужно принять тот факт, что персональные данные могут быть использованы различными способами и для разных целей, которые на сегодняшний день неочевидны. В связи с этим эксперты предлагают принять концепцию «разумных ожиданий». Это учет преимуществ, которые будут получены от создания систем ИИ и рисков при использовании персональных данных в неочевидных на данный момент целях, которые станут очевидными лишь с развитием технологий ИИ.

Эксперты считают необходимым указать в Декларации, что принятие концепции «разумных ожиданий» гарантирует, что создатели ИИ должны оценивать риски нецелевого использования персональных данных и преимуществ, которые даст ИИ технология конкретным индивидуумам, группам или обществу в целом.

Для контроля за ИИ технологиями предлагается ввести их государственную сертификацию. Отмечу, что сертификация для систем робототехники уже разрабатывается. К примеру Фонд ответственной робототехники (Foundation for Responsible Robotics)[6] предложил знак качества FRR, который гарантирует, что роботы и ИИ технологии созданы с уделением должного внимания правам человека и гуманитарным ценностям.

По мнению экспертов, ИИ должен помочь избежать человеческого фактора в принятии решений. Для недопущения дискриминации алгоритм ИИ должен быть протестирован через специальные категории персональных данных: гендерные данные, сведения о расовой, национальной принадлежностях, информация о здоровье. Отказ в доступе к таким данным приведет к тому, что ИИ будет выдавать необъяснимые дискриминационные решения. Поэтому следует применять риск-ориентированную модель и сравнивать преимущества использования ИИ технологий с рисками нарушения права человека на частную жизнь и конфиденциальность персональных данных.

В качестве примера приведу ИИ систему Fundus Machine, разработанную компанией Baidu в сотрудничестве с китайскими больницами. Обученная на большом количестве точно

промаркированных изображений глазного дна, она достигла диагностической точности, сопоставимой с результатами профессионального офтальмолога с опытом работы более 10 лет [7]. Точность диагностирования зависит от исходной информации (обучающего материала, представляющего собой сведения о здоровье), которую необходимо постоянно обновлять. Кроме того, существует риск системного сбоя и постановки неверного диагноза. Как заявляют разработчики Fundus Machine, использование ИИ-технологии связано с высокой степенью ответственности в отношении этики и безопасности, поэтому требуется, чтобы уровень стандартов проектирования и надежного ИИ был выше уровня, необходимого для любых новых технологий в прошлом.

Как обстоят дела в России?

10 октября 2019 г. Президент РФ подписал Указ № 490 «О развитии искусственного интеллекта в РФ» [8], которым одобрил Национальную стратегию развития искусственного интеллекта на период до 2030 года как часть национальной программы «Цифровая экономика РФ». Стратегия установила принципы развития и использования ИИ в России; в их число входит:

- защита прав и свобод человека, в том числе права на труд (!), и предоставление гражданам возможности получать знания и приобретать навыки для успешной адаптации к условиям цифровой экономики;
- безопасность — недопустимость использования ИИ технологий в целях умышленного причинения вреда гражданам и юридическим лицам;
- прозрачность — объяснимость работы ИИ технологий;
- технологический суверенитет — обеспечение необходимого уровня самостоятельности РФ в области ИИ посредством преимущественного использования отечественных технологий ИИ;

> разумная бережливость — принятие мер, направленных на реализацию государственной политики в научно--технической и других областях.

Российская Стратегия совпадает с европейским правовым регулированием только в принципе «прозрачности». Остальные принципы, если и схожи в формулировках, то различаются по сути.

В ЕС конфиденциальность персональных данных, составляющих базу для создания и обучения ИИ технологий, стоит во главе наряду с обеспечением прав и свобод человека и гражданина. Статья же 49 российской Стратегии гласит, что для создания ИИ нужны исходные данные, в том числе персональные, и государство обеспечит «благоприятные правовые условия (в том числе посредством создания экспериментального правового режима) для доступа к данным... в том числе собираемым государственными органами и медицинскими организациями».

То есть государство введет экспериментальный правовой режим, благодаря которому само определит, как распорядиться персональными данными своих граждан, и предоставит базы персональных данных разработчикам ИИ технологий, что позволит реализовать в полной мере технологический суверенитет РФ.

Если разумная бережливость в понимании экспертов ЕС — это благополучие общества и сохранение окружающей среды, то для России — это реализация государственной научно--технической политики. Если безопасность для ЕС — это предотвращение любого вреда, то в соответствии с российской Стратегией — это предупреждение лишь умышленного вреда.

Нет в Стратегии ни слова про отчетность, ответственность при создании ИИ, недискриминацию, справедливость и уважение самостоятельности человека, равно как и про этические правила надежного ИИ (есть только поручение Правительственной комиссии по цифровому развитию «разработать этические правила взаимодействия человека с ИИ»).

Вместо послесловия

Определить правовые рамки для регулирования ИИ можно. Именно сейчас происходит формирование правового отношения к ИИ технологиям и принципов создания и работы с ними. Однако следует учитывать, что эти принципы и их сущностное наполнение — это не просто ячейки, где можно поставить галочки и забыть. Это, пусть верхнеуровневое, но руководство к действию. Руководство к действию для тех, кто создает и работает с ИИ, а не для самих систем ИИ.

Это мы должны быть справедливыми, никого не дискриминировать и обеспечивать конфиденциальность, стремиться к социальному и экологическому благополучию. Это мы должны действовать этично и уважать права и свободы человека. Если мы будем следовать этим принципам, то и системы ИИ будут следовать за нами. От нас зависит, какой ИИ мы создадим и какие последствия наступят от его использования.

---

[1] High--Level Expert Group on Artificial Intelligence. — https://ec.europa.eu/digital--single-market/en/high-level--expertgroup--artificial-intelligence.

[2] Starship. — https://www.starship.xyz.

[3] Guidelines on Artificial intelligence and Data Protection — https://rm.coe.int/guidelines-on-artificial--intelligence-anddata-protection/168091f9d8.

[4] Declaration on Ethics and Data Protection in Artifical Intel ligence. — https://www.privacyconference2018.org/system/files/2018-10/20180922_ICDPPC_40th_AI-Declaration_AD_OPTED.pdf.

[5] https://www.informationpolicycentre.com/up_loads/5/7/1/0/57104281/cipl_response_to_icdppc_declara
.

[6] https://responsiblerobotics.org.

[7] Chinese search engine Baidu launches AI powered camera to detect eye fundus. — https://www.mobihealthnews.com/content/chinese--search-engine--baidu-launches-ai-powered-camera-de..

[8] Официальный интернет-портал правовой информации — http://pravo.gov.ru.tect-eye-fundus.

Наталья Ковалева Data protection officer ООО «Хэдхантер», канд. юрид. наук, CIPM

СТАТЬИ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ