

Уроки кросс-культурного менеджмента в преодолении предкризисной ситуации в этике в сфере искусственного интеллекта

С. С. Кладько

Автономная некоммерческая организация «НЕЙМАРК»

s2kladko@gmail.com

Аннотация

Этика в сфере искусственного интеллекта (ИИ), несмотря на действительную важность так называемого «мягкого права» для IT-сферы и номинально растущую популярность, сегодня уже столкнулась с целым рядом вызовов, позволяющим характеризовать текущую ситуацию в ней как предкризисную. Среди прочих вопросов особое внимание за последнее время стали привлекать культурные установки акторов, зачастую оказывающиеся дополнительным препятствием для применения этических принципов в практической деятельности. Наблюдая определённое сходство между существующим отношением различных аудиторий (прежде всего, разработчиков и владельцев ИИ-решений) к этике в сфере ИИ и отношением бизнеса к разработкам кросс-культурного менеджмента на ранней стадии его развития, автор предлагает экстраполировать применение определенных инструментов кросс-культурного менеджмента на вопросы формирования этической повестки дня для различных решений на основе искусственного интеллекта в различных отраслях с целью придания данным этическим принципам в сфере ИИ практической направленности.

Ключевые слова: этика в сфере искусственного интеллекта, кросс-культурный менеджмент, цифровая этика, искусственный интеллект, культурные измерения

Библиографическая ссылка: Кладько С. С., Уроки кросс-культурного менеджмента в преодолении предкризисной ситуации в этике в сфере искусственного интеллекта // Информационное общество: образование, наука, культура и технологии будущего. Выпуск 8 (Труды XXVII Международной объединенной научной конференции «Интернет и современное общество», IMS-2024, Санкт-Петербург, 24–26 июня 2024 г. Сборник научных статей). — СПб.: Университет ИТМО, 2024. С. 67–79. DOI: 10.17586/2587-8557-2024-8-67-79.

1. Введение

Текущее положение дел в этике в сфере искусственного интеллекта (ИИ), на первый взгляд, характеризуется положительной динамикой, о чём, в частности, может свидетельствовать почти пятикратное увеличение публикаций на эту тематику за последние пять лет [1], постоянный рост количества соответствующих деклараций и кодексов (а их количество сегодня составляет по разным источникам от 121 [2] до почти 200 [3]) и постоянно присоединяющихся к ним организаций, а также сохраняющийся уже долгое время интерес к этой теме государственных органов различных стран. Всё это, казалось, должно настраивать на оптимистичный лад в отношении того, что различные акторы (то есть «участники отношений в сфере искусственного интеллекта» [4]), наконец, осознали важность принятия во внимание этических вопросов и, соответственно, всё активнее следуют им на практике. Однако наравне с возрастающим интересом к этике в сфере ИИ приходят новые вопросы и новые разочарования. Так что сегодня, на самом деле, в пору

говорить о своеобразной предкризисной ситуации в этой сфере, о чём могут свидетельствовать следующие факты.

Смена вектора критики: от неопределённости к беспомощности. Несколько лет назад основные упреки в отношении этики в сфере ИИ сводились в основном к её неопределённости в плане функций, несправедливости по отношению к игрокам, не являющимися крупными и не владеющими дополнительными (в том числе административным) ресурсами, и даже к отсутствию единого определения [5]. Тем не менее, отношение к предмету обсуждения характеризовалось если не заинтересованностью, то определённым любопытством (хотя и с традиционным скепсисом, прежде всего со стороны разработчиков). Однако уже с 2022 г. постепенно усиливаются претензии, касающиеся полной практической бесполезности декларируемых этических принципов. Вряд ли это могло быть иначе, коль скоро разработка ИИ-решений происходит в этически пустой среде, которая не предполагает ни этического обучения в процессе образования, ни этического применения в IT-сфере [6]. Этика в сфере ИИ находится в противоречии с жёсткой конкуренцией при достаточно ограниченных ресурсах реального бизнеса, а сами этические принципы чаще всего не рассматриваются в целостности, но по отдельности [7], что мешает формированию общего видения, а, следовательно, и акторского консенсуса. Наконец, неразработанность механизма практической имплементации данных принципов, а также отсутствие единства в понимании профессиональной этики в IT-сфере (в отличие, например, от медицины, сфера высоких технологий куда более индивидуализирована) делают этику в сфере ИИ по-настоящему «беззубой» [8]. И это далеко не полный перечень претензий, число которых растёт куда динамичнее реальных кейсов позитивного следования этическим принципам.

Разрыв теории и практики и вынужденный «камуфляж». Огромное количество источников вряд ли гарантирует переход в качество, то есть к тому, чтобы этические принципы в сфере ИИ начали действительно соблюдаться если не всеми, то большинством акторов, которым справиться с постоянным ростом количества соответствующих инструкций, деклараций и кодексов оказывается зачастую не под силу. Более того, неудержимый прессинг растущего количества документов начинает охватывать и всё новые целевые аудитории, соответственно создавая новые вызовы, с которыми будет сталкиваться имплементация этих принципов на практике. Иными словами, происходит реальный разрыв между теоретическими обсуждениями и практикой их применения [9]. Такое «изобилие» только дополняет негативный настрой по поводу этики уже не только у разработчиков или владельцев ИИ-решений, но и, как указывалось выше, даже у представителей научного сообщества, всегда бывших активными сторонниками внедрения цифровой этики и начинающих испытывать сейчас определённое разочарование или даже усталость. С другой стороны, настойчивость государства в плане следования большому количеству не до конца определённых этических принципов в сфере ИИ заставляет акторов переходить к практике так называемого этического камуфляжа (bluewashing) [10]. Впрочем, по мнению некоторых исследователей, расплывчатость формулировок этических принципов на самом деле отлично используется индустрией как некая защита от куда более обязательных к исполнению норм правового регулирования [11]. И в том, и в другом случае этические принципы явно не используются по своему назначению и, по сути, решают вопросы имитации деятельности, что также можно отнести к традиционным характеристикам кризиса.

Невозможность «максимально этичного» синхронного использования этических принципов. Существует целый ряд исследований, в которых авторы делают оправданные и достаточно обоснованные попытки «сократить» по аналогии огромное число этических принципов, прописанных в различных документах [2, 3, 12]. Интересным здесь выглядит тот факт, что разница между средним количеством этических принципов, выдвигаемых практиками из IT-индустрии и научным сообществом, не такая и большая (9,1 против 10,2) [13], что может свидетельствовать об определённом консенсусе. В целом же, эксперты

предпочитают оперировать 5-7 характеристиками, что теоретически должно облегчить применение последних в реальной деятельности. Однако, даже с этим минимальным количеством возникают ситуации, при которых следование на практике одному этическому принципу неизменно осуществляется в ущерб другому [14], что ещё больше затрудняет желание разработчиков (да и, можно предположить, что и других акторов) соблюдать на практике эти принципы вообще. Иными словами, складывается парадоксальная ситуация, в которой акторы, пытаясь определить, соответствует ли то или иное ИИ-решение этическим принципам, могут понести некоторые временные издержки, а в результате придут к выводу, что соответствовать им в целом невозможно. Что и говорить, не лучшая ситуация, чтобы относиться к навязываемым, но до конца не проработанным принципам этики в сфере ИИ сколь-нибудь позитивно.

Этический универсализм vs культурные особенности. Речь идёт о недооценке влияния на процесс внедрения ИИ-решения фактора культурных ценностей, исторически сложившихся в конкретной целевой группе. В сфере IT контакт больших технологий и локальных культурных установок всегда представляет собой столкновение, в ходе которого даже самые общие этические принципы (например, «справедливость» или «конфиденциальность») совершенно по-разному «переводятся» различными культурами [15]. Для этики в сфере ИИ основной проблемой здесь является бездумный перенос этических констант глобального (чаще всего западного) характера на почву другой культуры, имеющей о содержании этих констант зачастую совершенно другое представление [16, с. 1096]. Как следствие, устоявшаяся культурная среда воспринимает универсализированные этические принципы как чуждые и даже враждебные, оказывая им на практике стабильное сопротивление [17]. Таким образом, недооценка культурных особенностей способна действительно осложнить применение положений этики в сфере ИИ, усугубляя в ней, тем самым, кризисные явления.

Следует отметить, что первые три вышеуказанные характеристики не являются чем-то новым: в той или иной степени, они были достаточно очевидны с самого начала роста интереса к этике в сфере ИИ, о них периодически упоминали и остаётся лишь удивляться тому, что к ним не отнеслись с должной серьёзностью раньше. А вот «культурный» фактор, ещё более усугубивший предкризисную ситуацию, привлёк к себе внимание не так давно, и пока по нему не предлагается каких-то решений, что сообщает его исследованию дополнительную актуальность и интерес в изучении.

2. Проблема

Несмотря на все существующие и потенциальные кризисы, цифровая этика, вне всякого сомнения, всегда останется важной составляющей общей регулятивной системы в сфере искусственного интеллекта. При всём продолжающемся росте нормативных правовых актов в разных странах мира, исключительно юридический подход («жёсткое право») никогда не сможет полностью охватить такую сложную и, самое главное, динамичную сферу с таким количеством разнородных акторов, которые вряд ли смогут строить консенсус, опираясь исключительно на правовое регулирование. IT-отрасль всегда славится постоянным «взрывным» развитием, за которым не сможет успеть ни одно законодательство, в связи с чем система этических принципов, выступая как элемент «мягкого» регулирования, действительно способна заполнить лакуны, возникающие в профессиональной коммуникации вследствие бурного роста индустрии. Это прекрасно осознают и правительственные органы различных государств, продолжая стимулировать распространение этических принципов среди максимально возможного большинства игроков национальных IT-рынков. Исходя из этого, основной вопрос стоит не в том, чтобы «отменить» этику в сфере ИИ как малоэффективный, забюрократизированный и не вносящий ясности инструмент, а в том, чтобы этика, приняв во внимание некоторые существенные моменты, наконец-то смогла показать свою максимальную полезность,

подтвердив самые высокие ожидания как государства, так и других акторов, или, иными словами, связать этические принципы ИИ с организационной практикой [18]; понять то, каким образом имея достаточно хорошо описанное «что» перейти к его практическому применению «как» [19]. Соответственно, общей исследовательской проблемой здесь является то, каким образом можно, наконец, сделать этические принципы в сфере ИИ хотя бы в общем применимыми и, главное, не мешающими взаимодействию акторов. Если же специфицировать проблему для данной работы, то автор видит её в том, что в настоящее время этика в сфере ИИ де-факто десинхронизирована с культурными установками акторов, что действительно мешает более эффективному и беспрепятственному внедрению соответствующих решений на практике.

3. Цель и вопросы

В данной статье предпринимается попытка определить варианты практического эффективного взаимодействия универсальных принципов этики в сфере ИИ и культурных особенностей различных акторов в рамках создания и дальнейшей эксплуатации решений на основе ИИ. Ввиду того, что на данный момент методы практического применения принципов этики в сфере ИИ носят спорадический характер на уровне отдельных кейсов, не связанных между собой в единую систему, следует обратить внимание на успешный опыт других сфер деятельности (так называемая адаптация по аналогии). Такая практика ориентирования на подходы, хорошо зарекомендовавшие себя в других отраслях деятельности, представляется вполне допустимой для этики в сфере ИИ, одним из примеров чего является выведение критериев цифровой этики по аналогии с принципами биоэтики [12] (да и вообще, учитывая всё большую «междисциплинарность» научной и практической деятельности, было бы странно хотя бы не посмотреть на опыт других отраслей). На взгляд автора, вариантом более эффективного преодоления препятствий, связанных с культурными установками при внедрении ИИ-решений, может стать более тщательное изучение подходов к практической работе с культурными особенностями соответствующих целевых аудиторий. При всех связанных с культурой исследовательских направлениях, наиболее практически ориентированные методы, как представляется, реализованы в парадигме кросс-культурного менеджмента (далее ККМ), основоположники которого обратили внимание на то, что несоответствие деловых максим представителей одной культуры ожиданиям коллег, представляющих другие культурные общности, оказывает негативное влияние на взаимодействие в международном бизнесе. Подобные ситуации приводили к многочисленным недопониманиям и конфликтам, которые результировались в финансовые и репутационные потери, которые иногда оказывались роковыми даже для крупных корпораций, решивших осуществлять глобальную экспансию. Как представляется, обращение к опыту ККМ для этики в сфере ИИ, было бы интересным не только с точки зрения преодоления культурных барьеров при внедрении и эксплуатации соответствующих ИИ-продуктов, но и с точки зрения определённого сходства в развитии этих двух направлений: совпадение временного старта их активизации (50-60-е годы XX века), надежды на то, что их применение будет способствовать улучшению существующей коммуникации, а также первоначальный огромный скепсис профессионального сообщества, не желавшего принимать во многом теоретические наработки к сведению (и опыт преодоления этого отношения в ККМ является действительно хорошей инструкцией для этики в сфере ИИ). Сегодня хотя бы общая осведомлённость о том, как использовать на практике инструменты кросс-культурного менеджмента, является необходимым условием ведения бизнеса в инокультурной среде. Принимая во внимание, что своеобразный «старт» двух направлений деятельности получился в чём-то похожим, можно предположить, что подходы, которые были использованы в ККМ, будут полезны и в том, чтобы помочь этике в сфере ИИ преодолеть существующий кризис.

Таким образом, основной целью данной статьи является идентификация подходов ККМ, которые могут быть применены в рамках более практической для акторов реализации положений этики в сфере ИИ. При этом представляется необходимым ответить на следующие вопросы:

- какие именно подходы из практики ККМ могут быть применены с целью сообщения дополнительного практического смысла принципам этики в сфере ИИ?
- на какие возможные вызовы при их применении следует обратить внимание?
- существуют ли ограничения по их применению?

4. Основные вызовы

Низведение этики исключительно до практики противоречит сути этики. Исходя из того, что в центре рассмотрения этики находятся мораль и нравственность, неизбежно возникает вопрос о том, а допустимы ли, например, для целей реальной индустрии практическая мораль и практическая нравственность? Скорее всего, подобные варианты звучат как некое лицемерие, подмена понятий, камуфляж (то есть то, с чем этике в сфере ИИ приходится зачастую сталкиваться сейчас). Соответственно, здесь существует большой риск того, что важные максимы, определяющие во многом бесконфликтное взаимодействие, приобретут утилитарные черты и, таким образом, утратят право определять моральное и аморальное (иными словами, будут чем угодно, но не этическими принципами).

Культурный субъективизм универсален и непреодолим. Согласно исследованиям, любые данные, применяемые при разработке ИИ-решений, обязательно находятся под влиянием культурных стереотипов акторов, участвовавших в сборе и переработке соответствующей информации, что является неизбежным и не подлежащим какой-либо коррекции [20]. Это неудивительно, так как невозможно отделить свои личные черты и убеждения от результатов своей работы (даже в области этики в сфере ИИ) [21]. Соответственно, этот вызов может лишить рассмотрение заявленной цели всякого смысла, так как может показаться, что для преодоления культурных барьеров этические принципы должны подстраиваться под каждый конкретный случай, а это нереально.

IT является крайне специфической сферой для социогуманитарных подходов. IT-сфера всегда отличается инновационностью, ускоренной динамикой развития, высочайшей конкуренцией и однозначным приоритетом STEM над HASS. Специалисты в данной отрасли вряд ли склонны акцентировать достаточное внимание на гуманитарные вопросы, требующие долгой рефлексии, а участие экспертов по этике в сфере ИИ в разработке ИИ-решений вызывает скептицизм как из-за ограниченного технологического опыта, так и из-за того, что и сами эксперты, говоря об этике, могут во многом не придерживаться таковой [21]. Возможно, что именно это объясняет ситуацию, когда всего лишь за несколько дней в марте 2022 г. ведущие западные IT-гиганты при всех красивых словах о необходимости обязательного учёта этических принципов полностью избавились от своих департаментов по этике в сфере ИИ, посчитав, что от их деятельности нет никакой реальной пользы [22]. Это действительно очень серьёзный вызов цели данной статьи, так как нужно найти аргументы, которые бы отвечали указанным ожиданиям хотя бы отчасти.

5. Предлагаемые подходы

К подходам, действующим в рамках ККМ и чьё применение в рамках этики в сфере ИИ могло бы быть полезно на практике, на взгляд автора, могут относиться следующие.

1. Формирование относительно понятной для участников (акторов) системы измерений. Этика в сфере ИИ и ККМ действуют в средах, мало предназначенных для точных измерений ввиду большой степени субъективизма и эмоциональности. Тем не менее, это не помешало ККМ за время своего существования сформировать значительное количество систем культурных измерений (начиная от «классических» Э. Холла,

Г. Хофстеде, Ф. Тромпенаарса и заканчивая современными вариациями Э. Мейер и Т. Джексона), которые сегодня достаточно широко используются в международной деловой практике. Конечно, вариантов подобных метрик очень много и каждый из них имеет свои существенные недостатки, их можно и нужно критиковать за некий культурно-географический детерминизм и за некоторую эмоциональность оценок, но, в целом, эти методы понятны и определяют очень важные критерии, на которые во многом следует обращать внимание при работе с категориальным аппаратом этики в сфере ИИ.

Следует отметить, что культурные измерения уже используются в ряде случаев сейчас для того, чтобы определиться с тем, каким образом местные культуры влияют на введение технологических (в том числе и ИИ) решений в разных регионах России [23], а практика их применения в работе с этикой в сфере ИИ уже имеет определённый опыт. Так, например, применение измерений Г. Хофстеде «индивидуализм-коллективизм» помогает разработчикам ИИ-решений в разработке систем, учитывающих этико-культурные особенности целевых аудиторий [24], а измерение «дистанция власти» применяется в исследованиях по приоритетам защиты беспилотных транспортных средств [25]. Также ряд культурных измерений применяется в исследованиях по выработке рекомендаций для разработчиков по этическому тестированию ИИ-решений [26].

Также вряд ли кто-то из акторов может безапелляционно настаивать на том, что в данном случае необходимо иметь 100% точные формулы для измерения этики в сфере ИИ (да и в ККМ): оценить социогуманитарный аспект настолько аккуратно вряд ли возможно, а попытки сформулировать «идеальные» метрики неизбежно приведут к так называемой «ловушке Гудхарта», что, по сути, лишь укрепит тенденцию этического «камуфляжа», но не поможет реальной практике. Однако отсутствие хотя бы общих метрик неизбежно влияет на коммуникацию акторов, трактующих тот или иной этический принцип исключительно с точки зрения индивидуального восприятия или корпоративной культуры. В условиях, когда государство постоянно напоминает о необходимости быть этичными в ИИ-решениях, а разработчики могут только строить догадки о том, что же на практике означает подобное требование от государственных органов, специально разработанные метрики, построенные, в том числе, на основе культурных измерений, позволили бы акторам достичь некоего единообразия в трактовках процессов и результатов, избежать тех же дискриминационных предвзятостей и, следовательно, достичь более эффективного взаимодействия.

2. Осознание и создание добавленной стоимости (added value). Концепции ККМ изначально были сориентированы на то, чтобы облегчить различные виды активностей в международном бизнесе, то есть всегда ориентировались на то, чтобы сообщить дополнительную эффективность в достижении сторонами, представляющими различные деловые культуры, так называемого взаимовыгодного решения. При этом ККМ не пытался занимать доминирующие позиции (до сих пор оставаясь лишь одним из направлений куда более широкой по объёму дисциплины под названием «Организационное поведение»), но также старался максимально абстрагироваться от наукообразия и чрезмерной теоретизации (хотя полностью избежать последней удавалось не всегда). Можно предположить, что подобная линия поведения во многом способствовала тому, что международный бизнес постепенно стал применять наработки в области ККМ, повышая доверие к ним. Как следствие, на сегодняшний день ККМ во многом продемонстрировал свою полезность в формировании added value для достижения результата.

Этика в сфере ИИ пока представляет собой некую непонятную абстрактную нагрузку, следование которой может (при сохранении нынешней ситуации) означать для тех же разработчиков временные и, как следствие, финансовые потери, что, в итоге, может способствовать утрате конкурентности продукта. И здесь существующие данные о том, что простое наличие политики этичного использования ИИ в организации уже существенно повышает степень доверия у большинства клиентов [27], выглядят не очень убедительными, так как здесь речь может просто идти об этическом камуфляже, о котором было упомянуто

ранее. Следует честно признать: современный бизнес справедливо не любит абстрактные идеи, а если готов воспринять что-то от научной среды, то основным условием здесь является возможность применения подобного предложения к практической пользе. Конечно, здесь можно высказывать сомнение, что даже при репутационных потерях этические механизмы слишком слабы, чтобы повлиять на состояние дел в бизнесе, однако имеющаяся на сегодняшний день практика показывает то, что этика в сфере ИИ вполне способна к этой роли, а несоблюдение этических принципов при разработке и работе с ИИ-решениями уже обрело критерии «монетизированности» [28]. Пожалуй, именно в таком случае этика в сфере ИИ перестаёт быть тяжелой и непонятной обязанностью для разработчика, но становится стимулом к тому, что быть этическим в работе с данными может быть полезным для бизнеса. Наконец, так же как недооценка культурного фактора заставляла даже ведущие мировые компании уходить с местных рынков, так и несовпадение декларируемых этических конструктов с их восприятием в местных культурах может послужить хотя и формальным, но достаточным поводом, чтобы при наличии даже отличного разработанного ИИ-решения требовать у разработчика серьёзно снизить цену на него. Исходя из сказанного, этика в сфере ИИ действительно может многому научить у ККМ в отношении того, как актерам, по крайней мере, обезопасить себя от не прямых финансовых потерь.

3. Более широкое внедрение фасилитации в диалоге акторов. Рассуждая о диалоге как условии эффективной коммуникационной деятельности, российский учёный Т. М. Дридзе сформулировала основные положения, при которых этот процесс будет успешен. К ним, в частности, относятся следующие:

- диалог есть нечто большее, чем обмен текстовыми сообщениями, представляя собой, прежде всего, контакт смыслов;
- важным условием диалога является стремление сторон к адекватной интерпретации коммуникативных намерений друг друга;
- в диалоге как в смысловом контакте очень важна роль посредника, помогающего сторонам находить правильные интерпретации, выходящие за исключительно лингвистические рамки [29].

На сегодняшний день с сожалением приходится констатировать фактически полное несоблюдение первого и второго положений в рамках взаимодействия акторов в сфере ИИ, а без этого вряд ли возможно позитивное восприятие декларируемых этических принципов. Подобная проблема в области международного бизнеса находится под постоянным интересом со стороны практиков в области ККМ, ведь здесь для совершения удачной сделки крайне важно то, чтобы стороны имели максимально точную интерпретацию смыслов. Именно поэтому ККМ уделяет столь много внимания практическим исследованиям в области диалога, накопив за несколько десятилетий своего существования обширнейший багаж в этой области (от урегулирования прямых конфликтов до прогнозирования потенциальных разногласий). Особое значение в этой деятельности имеет медиация (фасилитация), реализующая собой третье положение успешного диалога и помогающая избежать возможных ошибок в интерпретации культурных интенций партнёров, заинтересованных в сотрудничестве.

Если рассуждать о медиации в рамках этики в сфере ИИ, то наличие стороннего эксперта, обладающего необходимым опытом и знаниями и способного помочь актерам обнаружить культурные и этические противоречия при внедрении ИИ-проекта в жизнь, перекликается с одобрением этой практики российским Кодексом этики в сфере ИИ, приветствующем «проведение оценки рисков посредством привлечения нейтральной третьей стороны» [4]. Конечно, хотелось бы в этой связи говорить именно о медиаторах, то есть специально обученных и сертифицированных экспертах. Однако обеспечить подобного рода специалистами огромное, постоянно растущее количество IT-проектов представляется невероятным, да и акторы, как можно предположить, могут, исходя из положений Кодекса, трактовать каждый спорный случай как особый, неминуемо усложняя

процедуру согласования интересов. Именно поэтому для этики в сфере ИИ важна именно фасилитация, то есть открытый коммуникационный процесс, позволяющий сторонам или их представителям, обладающим хотя бы общими знаниями и навыками в области культурно-этического анализа, разрешать конфликтные вопросы с минимальными временными затратами. И здесь успешный опыт ККМ поистине огромен.

6. Ограничения

Следует признать, что предлагаемые варианты того, что именно этика в сфере ИИ может взять у ККМ с целью преодоления существующего кризиса, не лишены ограничений и даже противоречий. Прежде всего, применение опыта тех же культурных измерений с учётом огромного количества вариантов и методик может и само оказаться достаточно затратным по времени, оказывая, тем самым, не менее ограничивающее влияние на внедрение или эксплуатацию ИИ-решения. Порою, чтобы разобраться в хитросплетениях всего количества измерений, может потребоваться отдельный ознакомительный курс, но и тогда неизбежно встанут вопросы о временных затратах на проведение всестороннего анализа вводимых данных ещё и методами ККМ (что противоречит задаче по недопущению осложнения работы тем же разработчикам). Методом преодоления этого ограничения, может быть, выбор определённых измерений (не более двух-трёх), а также понимание того, где именно можно будет найти уже готовую информацию. Благо такие описания от экспертов находятся в открытом доступе, что позволяет сделать своеобразный экспресс-анализ, результатов которого может быть достаточно для понимания того, какие этико-культурные проблемы могут возникнуть при использовании продукта. Следующим ограничением может выступать недостаточный набор практических примеров того, каким образом методы ККМ оказываются полезными для этики в сфере ИИ. На сегодняшний день есть достаточное количество работ, которые описывают такую возможность в теории, однако из практики автору пока известны только примеры, которые были приведены в работе. Безусловно, данное положение нуждается в дополнительном тестировании по мере появления конкретных данных (пока же массив информации носит, скорее, агитационно-умозрительный характер). Наконец, дополнительным ограничением может выглядеть и сам характер ККМ, наработки которого хотя и получают всё большее количество положительных отзывов, но до сих пор не являются однозначно применяемым методом в международном бизнесе. Однако, даже имеющиеся результаты того, как социогуманитарное направление «завоевало» право эффективно помогать на практике, заслуживают изучения такого опыта для этики в сфере ИИ.

В то же время, вряд ли можно согласиться с тем, что ещё одним ограничением может являться построение рассуждений о взаимодействии этики в сфере ИИ и ККМ на основе англоязычных источников. Во-первых, на сегодняшний день подавляющее число работ, рассматривающих проблематику этики в сфере ИИ в рамках её взаимодействия с культурными установками соответствующих целевых аудиторий, пока ещё принадлежат западным специалистам. Во-вторых, речь идёт не о смысловом наполнении того или иного этического принципа, а об общем анализе применимости этики в сфере ИИ, а здесь подходы достаточно универсальны. В любом случае, следует согласиться с тем, что этика в сфере ИИ ещё очень молодая область, чтобы успешно инкорпорировать ценности различных культур, и что она делает лишь первые шаги, чтобы перестать в своих принципах быть ориентированной в основном на западные традиции [30].

Следует также добавить, что применение рекомендуемых подходов ограничивается ИИ-решениями, предназначенными исключительно для гражданского использования.

7. Заключение

Роль этики в сфере ИИ, несмотря на активизировавшееся нормотворчество для ИИ-ойкумены, вряд ли станет менее значимой в будущем ввиду характерных особенностей ИТ-индустрии, быстрое развитие которой не оставляет шансов правовой системе обеспечить здесь всестороннее регулирование. При этом тесная историческая связь культуры и этики однозначно делает культурный фактор не просто дополнительным, но особо значимым на любом этапе работы с ИИ-решением (от сбора и переработки данных до непосредственной работы продукта) во избежание непредвиденных временных, репутационных и, возможно, финансовых потерь. Можно, конечно, возразить, что тот же ККМ сегодня всё чаще говорит о некоей универсальной, не знающей границ бизнес-культуре, а универсальным трактовкам принципов этики в сфере ИИ уже готовится некая идеальная среда, формируемая под влиянием так называемой «алгокогнитивной культуры», представляющей собой результат конвергентного взаимодействия человека и высокотехнологических решений в рамках цифровой гипер-сети, в которой функции традиционной культуры (передача знаний и социализация) переместились в виртуальное пространство [31]. Пожалуй, здесь действительно прекрасные условия, чтобы следовать единым пониманиям того, что означает тот или иной этический принцип, а в случае несогласия подвергнуться цифровой «отмене», чему уже есть целый ряд примеров. Однако, даже соглашаясь с тем, что цифровая культура (особенно в молодёжной среде) становится всё более универсальной, реальная культура с её важностью для самоидентификации индивида и социальной группы вряд ли утратит свою историческую роль под влиянием высоких технологий (особенно учитывая то, сколько раз в истории человечества проводились подобные безрезультатные эксперименты над культурой). Да и государство не позволит, чтобы внешняя культурная зависимость подрывала основы государственности. Следовательно, этике в сфере ИИ для того, чтобы стать действительно применимой, следует не просто принимать во внимание культурные особенности акторов, но и учиться этому у наиболее удачных культурных практик, включая, среди прочих, и методы ККМ.

Литература

- [1] Artificial Intelligence Index Report. Stanford Institute for Human-Centered Artificial Intelligence (апрель 2023). URL: <https://aiindex.stanford.edu/report/> (дата обращения 10.01.2024).
- [2] Tidjon L. N., F. Khomh F. The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis // *IEEE Transactions on Artificial Intelligence*. 2023. Vol. 4, № 4. P. 820–839. DOI: 10.1109/TAI.2022.3225132.
- [3] Corrêa N. K., Galvão C. Santos J. W., et al. Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance // *Patterns* (N Y). 2023. Vol. 4 (10). 100857. DOI: 10.1016/j.patter.2023.100857.
- [4] Кодекс этики в сфере искусственного интеллекта РФ // Альянс в сфере искусственного интеллекта, 2021. URL: <https://ethics.a-ai.ru/> (дата обращения: 10.02.2024).
- [5] Pew Research Center. Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm Within the Next Decade (Report). 16 июня 2021. URL: <https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/> (дата обращения: 05.01.2024).
- [6] Munn L. The uselessness of AI ethics // *AI Ethics*. 2023. Vol. 3. P. 869–877. DOI: 10.1007/s43681-022-00209-w.
- [7] Lauer D. You cannot have AI ethics without ethics // *AI Ethics*. 2021. Vol. 1. P. 21–25. DOI: 10.1007/s43681-020-00013-4.
- [8] Green B. The Contestation of Tech Ethics: A Sociotechnical Approach to Ethics and Technology in Action // *Journal of Social Computing*. 2021. Vol. 2 (3). P. 209–225. DOI: 10.23919/JSC.2021.0018.

- [9] Bleher H., Braun M. Reflections on Putting AI Ethics into Practice: How Three AI Ethics Approaches Conceptualize Theory and Practice // *Science and Engineering Ethics*. 2023. Vol. 29 (3). P. 21. DOI: 10.1007/s11948-023-00443-3.
- [10] Floridi L. Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical // *Philosophy & Technology*. 2019. Vol. 32. P. 185–193. DOI: 10.1007/s13347-019-00354-x.
- [11] Mittelstadt B. Principles alone cannot guarantee ethical AI // *Nature Machine Intelligence*. 2019. Vol. 1. P. 501–507. DOI: 10.1038/s42256-019-0114-4.
- [12] Floridi L., Cowls J., Beltrametti M. et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations // *Minds & Machines*. 2018. Vol. 28. P. 689–707. DOI: 10.1007/s11023-018-9482-5.
- [13] Hagendorff T. The Ethics of AI Ethics: An Evaluation of Guidelines // *Minds & Machines*. 2020. Vol. 30. P. 99–120. <https://doi.org/10.1007/s11023-020-09517-8>.
- [14] Sanderson C., Douglas D., Lu Q. Implementing Responsible AI: Tensions and Trade-Offs Between Ethics Aspects // *International Joint Conference on Neural Networks (IJCNN)*, Gold Coast, Australia. 2023. P. 1–7. DOI: 10.1109/IJCNN54540.2023.10191274.
- [15] Hagerty A., Rubinov I. Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence // *ArXiv*. 2019. URL: <https://arxiv.org/abs/1907.07892> (дата обращения: 17.02.2024).
- [16] Roche C., Wall P.J., Lewis D. Ethics and diversity in artificial intelligence policies, strategies and initiatives // *AI Ethics*. 2023. Vol. 3. P. 1095–1115. DOI: 10.1007/s43681-022-00218-9.
- [17] Goffi E., Colin L., Belouali S. Ethical assessment of AI cannot ignore cultural pluralism: a call for broader perspective on AI ethics // *The International Journal of Human Rights*. 2021. Vol. 1 (2). P. 151–175.
- [18] Krijger J., Thuis T., de Ruiter M. et al. The AI ethics maturity model: a holistic approach to advancing ethical data science in organizations // *AI Ethics*. 2023. Vol. 3. P. 355–367. DOI: 10.1007/s43681-022-00228-7.
- [19] Morley J., Floridi L., Kinsey L. et al. From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices // *Science and Engineering Ethics*. 2020. Vol. 26. P. 2141–2168. DOI: 10.1007/s11948-019-00165-5.
- [20] Caliskan A., Bryson J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases // *Science*. 2017. Vol. 356 (6334). P.183–186.
- [21] Hagendorff T. AI ethics and its pitfalls: not living up to its own standards? // *AI Ethics*. 2023. Vol. 3. P. 329–336. DOI: 10.1007/s43681-022-00173-5.
- [22] Criddle C., Murgia M. Big tech companies cut AI ethics staff, raising safety concerns // *Financial Times*. 29.03.2023. URL: <https://www.ft.com/content/26372287-6fb3-457b-9e9c-f722027f36b3> (дата обращения: 20.02.2023).
- [23] Аузан А. А., Авдиенкова М. А., Андреева Д. А. и др. Социокультурные факторы инновационного развития и успешной имплементации реформ. Центр Стратегических разработок «Истина», 2017. 197 с.
- [24] de Wever E. Navigating the Ethical Landscape of Artificial Intelligence: A Cultural Perspective. September 27, 2023. URL: <https://labelnone.com/labelnone-marketing-blog/navigating-ethical-landscape-of-ai-a-cultural-perspective/> (дата обращения: 15.02.2024).
- [25] Liu Y. Cross-Cultural Challenges to Artificial Intelligence Ethics // *Computer Sciences & Mathematics Forum*. 2023. Vol. 8 (1). P. 21. DOI: 10.3390/cmsf2023008021.
- [26] Kladko S. Cultural-Ethical Evaluation in the Launch of AI Education Technologies // *Agents and Multi-agent Systems: Technologies and Applications*. 2023. Vol 354. P. 265–275. DOI: 10.1007/978-981-99-3068-5_24.
- [27] Farzan S. Ethics First: The Imperative of Responsible AI Adoption in Marketing // *Forbes*. 29.09.2023. URL: <https://www.forbes.com/sites/sunshinefarzan/2023/09/29/ethics-first-the->

- imperative-of-responsible-ai-adoption-in-marketing/?sh=2add0a9a1486 (дата обращения: 25.02.2024).
- [28] DataRobot's State of AI Bias Report Reveals 81% of Technology Leaders Want Government Regulation of AI Bias (Report). DataRobot. 2022. URL: <https://www.datarobot.com/newsroom/press/datarobots-state-of-ai-bias-report-reveals-81-of-technology-leaders-want-government-regulation-of-ai-bias/> (дата обращения: 30.01.2024).
- [29] Дридзе Т. М. Экоантропоцентрическая модель социального познания как путь к преодолению парадигмального кризиса в социологии // Социс. 2000. № 2. С. 20–28. URL: <https://www.isras.ru/files/File/Socis/02-2000/003.DRIDZE.pdf> (дата обращения: 20.02.2004).
- [30] Hickok M. Lessons learned from AI ethics principles for future actions // AI Ethics. 2021. Vol. 1. P. 41–47. DOI: 10.1007/s43681-020-00008-1.
- [31] Карелов С. Алгокогнитивная культура. Новый этап эволюции Homo sapiens // Medium.com. 22.03.2021. URL: <https://sergey-57776.medium.com/алгокогнитивная-культура-510dbadd149c> (дата обращения: 01.03.2024).

Lessons from Cross-Cultural Management in Overcoming the Pre-Crisis Situation in AI Ethics

S. S. Kladko

ANO «NEIMARK»

AI Ethics, despite the actual importance of the so-called 'soft law' for the IT sphere and its nominally growing popularity, is facing today a number of challenges that allow to characterize the current situation in it as pre-crisis. Among other issues the AI experts have to deal with nowadays are the cultural attitudes of actors. Initially not very noticeable, today the resistance of cultures to the spread of universal (basically Western) ethical principles, which turns out an additional obstacle to the development and future functioning of AI-based solutions. Observing certain similarities between the existing attitudes of various audiences (primarily, developers and owners of AI solutions) towards AI ethics and the attitudes of businesses towards Cross-Cultural Management at the early stage of their development, the author, on the basis of the source analysis, extrapolates the application of certain tools from Cross-Cultural Management to the issues of forming an ethical agenda for AI-based solutions in various industries and proposes the ways how such an approach can provide these ethical principles with a new practical boost.

Keywords: AI Ethics, Artificial Intelligence, Cross-Cultural Management, cultural dimensions, digital ethics

Reference for citation: Kladko S. S. Lessons from Cross-Cultural Management in Overcoming the Pre-Crisis Situation in AI Ethics // Information Society: Education, Science, Culture and Technology of Future. Vol. 8 (Proceedings of the XXVII International Joint Scientific Conference «Internet and Modern Society», IMS-2024, St. Petersburg, June 24–26, 2024). — St. Petersburg: ITMO University, 2024. P. 67–79. DOI: 10.17586/2587-8557-2024-8-67-79.

Reference

- [1] Artificial Intelligence Index Report. Stanford Institute for Human-Centered Artificial Intelligence (April, 2023). URL: <https://aiindex.stanford.edu/report/> (accessed date: 10.01.2024).

- [2] Tidjon L. N., F. Khomh F. The Different Faces of AI Ethics Across the World: A Principle-to-Practice Gap Analysis // *IEEE Transactions on Artificial Intelligence*. 2023. Vol. 4, № 4. P. 820–839. DOI: 10.1109/TAI.2022.3225132.
- [3] Corrêa N. K., Galvão C. Santos J. W., et al. Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance // *Patterns* (N Y). 2023. Vol. 4(10). 100857. DOI: 10.1016/j.patter.2023.100857.
- [4] Kodeks Etik v Sfere Iskustvennogo Intellecta [AI Code of Ethics of the Russian Federation] // Alians v sfereiskustvennogo intellekta [AI Alliance], 2021. URL: <https://ethics.a-ai.ru/> (accessed date: 10.02.2024). (In Russian)
- [5] Pew Research Center. Experts Doubt Ethical AI Design Will Be Broadly Adopted as the Norm Within the Next Decade (Report). June, 16. 2021. URL: <https://www.pewresearch.org/internet/2021/06/16/experts-doubt-ethical-ai-design-will-be-broadly-adopted-as-the-norm-within-the-next-decade/> (accessed date: 05.01.2024).
- [6] Munn L. The uselessness of AI ethics // *AI Ethics*. 2023. Vol. 3. P. 869–877. DOI: 10.1007/s43681-022-00209-w.
- [7] Lauer D. You cannot have AI ethics without ethics // *AI Ethics*. 2021. Vol. 1. P. 21–25. DOI: 10.1007/s43681-020-00013-4.
- [8] Green B. The Contestation of Tech Ethics: A Sociotechnical Approach to Ethics and Technology in Action // *Journal of Social Computing*. 2021. Vol. 2 (3). P. 209–225. DOI: 10.23919/JSC.2021.0018.
- [9] Bleher H., Braun M. Reflections on Putting AI Ethics into Practice: How Three AI Ethics Approaches Conceptualize Theory and Practice // *Science and Engineering Ethics*. 2023. Vol. 29 (3). 21. DOI: 10.1007/s11948-023-00443-3.
- [10] Floridi L. Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical // *Philosophy & Technology*. 2019. Vol. 32. P. 185–193. DOI: 10.1007/s13347-019-00354-x.
- [11] Mittelstadt B. Principles alone cannot guarantee ethical AI // *Nature Machine Intelligence*. 2019. Vol. 1. P. 501–507. DOI: 10.1038/s42256-019-0114-4.
- [12] Floridi L., Cowlis J., Beltrametti M. et al. AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations // *Minds & Machines*. 2018. Vol. 28. P. 689–707. DOI: 10.1007/s11023-018-9482-5.
- [13] Hagendorff T. The Ethics of AI Ethics: An Evaluation of Guidelines // *Minds & Machines*. 2020. Vol. 30. P. 99–120. <https://doi.org/10.1007/s11023-020-09517-8>.
- [14] Sanderson C., Douglas D., Lu Q. Implementing Responsible AI: Tensions and Trade-Offs Between Ethics Aspects // *International Joint Conference on Neural Networks (IJCNN)*, Gold Coast, Australia. 2023. P. 1–7. DOI: 10.1109/IJCNN54540.2023.10191274.
- [15] Hagerty A., Rubinov I. Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence // *ArXiv*. 2019. URL: <https://arxiv.org/abs/1907.07892> (дата обращения: 17.02.2024).
- [16] Roche C., Wall P.J., Lewis D. Ethics and diversity in artificial intelligence policies, strategies and initiatives // *AI Ethics*. 2023. Vol. 3. P. 1095–1115. DOI: 10.1007/s43681-022-00218-9.
- [17] Goffi E., Colin L., Belouali S. Ethical assessment of AI cannot ignore cultural pluralism: a call for broader perspective on AI ethics // *The International Journal of Human Rights*. 2021. Vol. 1 (2). P. 151–175.
- [18] Krijger J., Thuis T., de Ruiter M. et al. The AI ethics maturity model: a holistic approach to advancing ethical data science in organizations // *AI Ethics*. 2023. Vol. 3. P. 355–367. DOI: 10.1007/s43681-022-00228-7.
- [19] Morley J., Floridi L., Kinsey L. et al. From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices // *Science and Engineering Ethics*. 2020. Vol. 26. P. 2141–2168. DOI: 10.1007/s11948-019-00165-5.

- [20] Caliskan A., Bryson J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases // *Science*. 2017. Vol. 356 (6334). P.183–186.
- [21] Hagendorff T. AI ethics and its pitfalls: not living up to its own standards? // *AI Ethics*. 2023. Vol. 3. P. 329–336. DOI: 10.1007/s43681-022-00173-5.
- [22] Criddle C., Murgia M. Big tech companies cut AI ethics staff, raising safety concerns // *Financial Times*. 29.03.2023. URL: <https://www.ft.com/content/26372287-6fb3-457b-9e9c-f722027f36b3> (accessed date: 20.02.2023).
- [23] Auzan A. A., Avdienkova M. A., Andreeva D. A. Sotsiokul'turnye faktory innovatsionnogo razvitiya i uspešnoi implementatsii reform [Sociocultural Factors of Innovational Development and Successful Implementation of Reforms]. Tsentr Strategicheskikh razrabotok «IstinA» [Centre for Strategic Research «IstinA»], 2017. 197 p. (In Russian)
- [24] de Wever E. Navigating the Ethical Landscape of Artificial Intelligence: A Cultural Perspective. September 27, 2023. URL: <https://labelnone.com/labelnone-marketing-blog/navigating-ethical-landscape-of-ai-a-cultural-perspective/> (accessed date: 15.02.2024).
- [25] Liu Y. Cross-Cultural Challenges to Artificial Intelligence Ethics // *Computer Sciences & Mathematics Forum*. 2023. Vol. 8 (1). 21. DOI: 10.3390/cmsf2023008021.
- [26] Kladko S. Cultural-Ethical Evaluation in the Launch of AI Education Technologies // *Agents and Multi-agent Systems: Technologies and Applications*. 2023. Vol 354. P. 265–275. DOI: 10.1007/978-981-99-3068-5_24.
- [27] Farzan S. Ethics First: The Imperative of Responsible AI Adoption in Marketing // *Forbes*. 29.09.2023. URL: <https://www.forbes.com/sites/sunshinefarzan/2023/09/29/ethics-first-the-imperative-of-responsible-ai-adoption-in-marketing/?sh=2add0a9a1486> (accessed date: 25.02.2024).
- [28] DataRobot's State of AI Bias Report Reveals 81% of Technology Leaders Want Government Regulation of AI Bias (Report). DataRobot. 2022. URL: <https://www.datarobot.com/newsroom/press/datarobots-state-of-ai-bias-report-reveals-81-of-technology-leaders-want-government-regulation-of-ai-bias/> (accessed date: 30.01.2024).
- [29] Dridze T. M. Ekoantropocentricheskaya model' social'nogo poznaniya kak put' k preodoleniyu paradigmalnogo krizisa v sociologii // *Socis*. 2000. № 2. S. 20–28. URL: <https://www.isras.ru/files/File/Socis/02-2000/003.DRIDZE.pdf> (access date: 20.02.2004). (In Russian)
- [30] Hickok M. Lessons learned from AI ethics principles for future actions // *AI Ethics*. 2021. Vol. 1. P. 41–47. DOI: 10.1007/s43681-020-00008-1.
- [31] Karelov S. Algotivnaya kul'tura. Novyj etap evolyucii Homo sapiens // *Medium.com*. 22.03.2021. URL: <https://sergey-57776.medium.com/algokognitivnaya-kul'tura-510dbadd149c> (accessed date: 01.03.2024). (In Russian)