

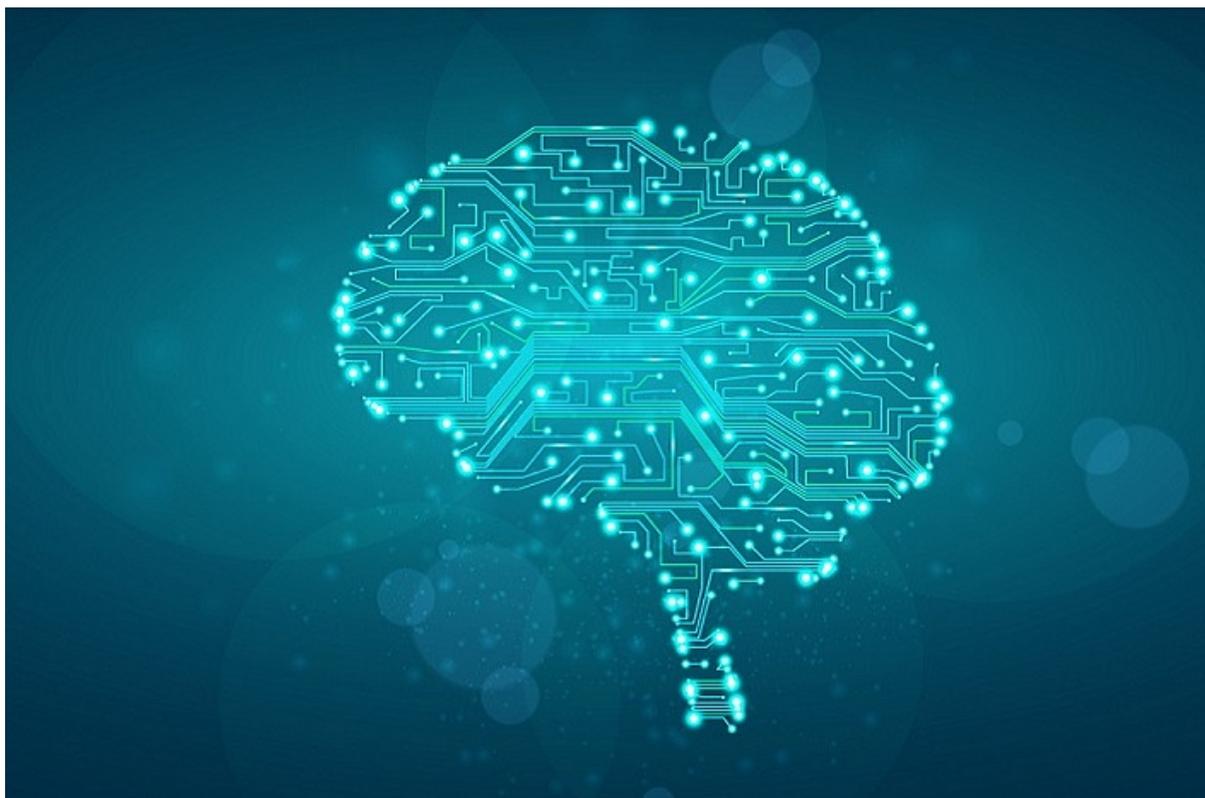
ИНДЕКС БЕЗОПАСНОСТИ

ПИР-ЦЕНТР | ТРИАЛОГ

№ 3 (7) | 2020

«Этичный искусственный интеллект» становится инструментом политики

Александр Федоров



В статье анализируются стратегии развития искусственного интеллекта (ИИ). Рассматриваются особенности подхода Европейского Союза к созданию этичного ИИ. Особое внимание автор уделяет деятельности Групп экспертов высокого уровня, ключевым принципам и требованиям, предъявляемым к системам ИИ, а также роли оборонно-промышленного комплекса в их разработке.

ГЛАВНОЕ



- **Более 30 стран и государственных объединений в той или иной форме уже определили свои национальные стратегии в области искусственного интеллекта (ИИ).** Однако у большинства из них работа пока ведется в рамках теоретических проработок или оформления начальных организационных решений.
- **По оценкам социологов, перспективы появления ИИ связываются не только с большими надеждами, но и со значительными опасениями.** У обывателя ИИ ассоциируются, в первую очередь, с кибер-солдатами и очередями безработных, у интеллигенции – с оруэлловским «большим братом».
- **Сегодня в Брюсселе разработка систем на основе ИИ воспринимается как естественная доминанта экономического развития.** 25 стран ЕС в апреле 2018 г. подписали декларацию о сотрудничестве в сфере ИИ (Declaration of Cooperation on Artificial Intelligence). На международном рынке уже идет конкурентная борьба производителей ИИ-продукции.
- **Если вопрос о том, существует ли рынок ИИ, для многих пока открыт, то потребность в таком товаре на рынке уже сформирована.** И большую помощь в этом оказали военные.

[Читать доклад на сайте ПИР-Центра](#)

ОБ АВТОРЕ

АЛЕКСАНДР ВАЛЕНТИНОВИЧ ФЕДОРОВ

Канд. физ.-мат. наук, старший научный сотрудник, доцент МГТУ им. Н.Э. Баумана, член Экспертного совета ПИР-Центра.

Окончил Московский государственный университет им. М.В. Ломоносова (факультет вычислительной математики и кибернетики) и аспирантуру МГУ. Прошел обучение на факультете повышения квалификации Дипломатической академии МИД РФ.

После окончания учебы работал в учреждениях государственного аппарата. Занимался проблемами разоружения и нераспространения ОМУ. Последние 25 лет ведет активную научную и практическую деятельность в сфере международной информационной безопасности и противодействия высокотехнологичному терроризму. Входил в состав российских делегаций на многих международных форумах по этой проблематике в форматах ООН, «G-8», Совета Европы, БРИКС, ШОС и др.

Входил в состав экспертного совета Комитета по безопасности Государственной думы 4-го созыва. Автор около сотни публикаций в различных научных и политических изданиях.

О СЕРИИ

Серия «Индекс Безопасности» – доклады, аналитические статьи, комментарии и интервью, которые отражают позиции российских и зарубежных экспертов по актуальным вызовам глобальной безопасности и политики России в этой сфере. Задача серии – дать понятный анализ проблем международной безопасности и предложить для них конкретные и реалистичные решения. Серия пришла на смену журналу «Индекс Безопасности», издаваемому ПИР-Центром в 1994 – 2016 гг. Авторы и редакторы серии будут рады комментариям, вопросам и предложениям, которые читатели могут направить на электронную почту inform@pircenter.org.



Occasional paper series #3 (7): Редакторы: *Вадим Козюлин, Надежда Кулибаба*

Наш электронный адрес: inform@pircenter.org

Здесь вы можете подписаться на другие рассылки ПИР-Центра или отписаться от них.

This email was sent to [<<Email>>](#)

[why did I get this?](#) [unsubscribe from this list](#) [update subscription preferences](#)

PIR Center · PO Box 147 · Moscow 119019 · Russia

«Этичный искусственный интеллект» становится инструментом политики

Александр Федоров



Европейский подход к ИИ

Более 30 стран и государственных объединений в той или иной форме уже определили свои национальные стратегии в области ИИ[i]. Однако у большинства

из них работа пока ведется в рамках теоретических проработок или оформления начальных организационных решений. К этой группе следует отнести и страны Европы, включая наиболее развитые, и США, многими, на самом деле, ошибочно считаемые безусловным лидером в области ИИ: по оценке заместителя министра обороны США по исследованиям и разработкам Майкла Гриффина, Соединенные Штаты «больше не обладают неоспоримым техническим превосходством [в сфере ИИ – Авт.], позволяющим доминировать в грядущей схватке» [ii].

Наиболее интересным и продвинутым в этом контексте следует, вероятно, рассматривать подход Евросоюза и его исполнительного органа Еврокомиссии. Итоги будут подводиться позже, но оценить европейское видение проблемы и принятые решения можно уже сейчас, и их анализ может быть, по мнению автора, весьма полезен при планировании и исполнении мероприятий в рамках реализации российской стратегии. Возможно, у читателя возникнет чувство того, что автор необоснованно выделяет Европу, лишь оговорившись о достижениях других стран, в частности США и Китая. Это не так. Просто, то, что мы ниже будем говорить о Европе, во многом показательно, происходит сегодня, сжато во времени и с учетом опыта тех же Вашингтона и Пекина, как, впрочем, и других стран.

В основе продвигаемого Еврокомиссией (ЕК) проекта лежит провозглашенная Урсулой фон дер Ляйен идея достижения Евросоюзом мирового лидерства в производстве систем искусственного интеллекта[iii]. При избрании фрау Ляйен новым президентом ЕК она в своей речи перед Европарламентом, воспринимаемой не иначе как политическая повестка дня нового европейского кабинета министров, назвала «искусственный интеллект, сделанный в Европе» целью цифровой стратегии ЕС. Основным аргументом Ляйен стал тезис о том, что Европа способна производить продукты ИИ более этичные, чем любые другие, созданные где-либо в мире[iv].

По оценкам социологов, перспективы появления ИИ связываются не только с большими надеждами, но и со значительными опасениями. У обывателя ИИ ассоциируются, в первую очередь, с кибер-солдатами и очередями безработных, у интеллигенции – с оруэлловским «большим братом».

«Мифы вокруг искусственного интеллекта связаны в первую очередь с тем, что искусственный интеллект — это что-то такое сложное,— как-то поделился своими и чужими фобиями Герман Греф,— недоступное, что это технология будущего. Второй миф: искусственный интеллект — это что-то связанное с роботами, которые наступают и которые пока еще сидят за углом или в засаде, но на самом деле целются на наши рабочие места и скоро нас заменят»[v]. Только не ясно, последнее это все-таки миф или прогноз.

В обыденном сознании мы наделяем «братьев наших меньших» человеческим разумом и переносим на них отношения права и морали человеческого общества (причем именно того, в котором выросли) и ждем, что любимая кошка не будет есть любимого хомячка. По такому же типу мы хотели бы строить отношения и миром ИИ, ожидая от него в ответ соблюдения, сформулированного Айзеком Азимовым нравственного императива «робот не может причинить вред человеку».

Но это обыденное сознание. А что подвигло Евросоюз пойти этим путем и выступить с инициативой выработать и предложить человечеству этические требования к ИИ, не раскрывая и даже не зная по сути, что это такое?

Этап дискуссий на тему нужности ИИ Европа в целом уже прошла. Сегодня в Брюсселе разработка систем на их основе воспринимается как естественная доминанта экономического развития. 25 стран ЕС в апреле 2018 г. подписали декларацию о сотрудничестве в сфере ИИ (Declaration of Cooperation on Artificial Intelligence)[vi]. На международном рынке уже идет конкурентная борьба производителей ИИ-продукции. В Париже даже выражают готовность пойти на двусторонние контакты с Берлином с тем, чтобы совместно одержать победу над такими ключевыми игроками как США, Россия и Китай. Франкенштейн остался в так любимых европейцами сказках, но Терминатор еще пугает с экранов перспективой наводнения мира «роботами-убийцами», наделенными недюжинным ИИ. Аналитиками Римского клуба развитие «недружественного ИИ» отнесено к сфере глобальных угроз человечеству[vii]. В этих условиях разработки «этичного ИИ», снимающие угрозу и обращенные к человеку, не только вызывают общественный интерес, но обладают инвестиционной и коммерческой привлекательностью. И вот в своей «тронной речи» Урсула фон дер Ляйен подчеркивает не просто особое значение для Европы производства своих систем ИИ, а делает акцент на разработке некоего «этичного ИИ», который якобы станет основой достижения Старым светом мирового лидерства. «За мои первые 100 дней пребывания в должности я буду продвигать законодательство, способствующее выработке скоординированного европейского подхода к человеческим и этическим последствиям искусственного интеллекта... Я позабочусь о том, чтобы мы отдавали приоритет инвестициям в искусственный интеллект, как через многолетнюю финансовую структуру, так и за счет более широкого использования государственно-частных партнерств»[viii].

Мотив моральности ИИ услышан и взят на вооружение и странами, не связанными общими с Евросоюзом интересами. В своем выступлении на 74-й сессии ГА ООН, премьер-министр Великобритании Борис Джонсон заявил, что самый этичный ИИ будет создан на Британских островах. Но, в отличие от других игроков, Еврокомиссия уже приступила к выработке практических мер. Одобрив в апреле 2018 г. «Сообщение об искусственном интеллекте» – документ, излагающий подход ЕС к ИИ, – ЕК нацелилась на разработку его этической и правовой базы с тем, чтобы на ее основе подготовить европейцев к социально-экономическим изменениям, вызванным внедрением ИИ в жизнь.

Деятельность экспертных групп

С этой целью в декабре 2018 г. под эгидой ЕК начала работу Группа экспертов высокого уровня (ГЭВУ) по ИИ (High-Level Expert Group on Artificial Intelligence) – 52 прошедших персональный отбор специалиста разного профиля из стран ЕС[ix]. Мандат ГЭВУ предусматривает подготовку для ЕК предложений и справок по этическим, правовым и социальным вопросам ИИ, включая его влияние на социально-экономическую сферу, а также по вопросам выработки мер реагирования на средне- и долгосрочные вызовы в сфере ИИ, готовить рекомендации по политике в этой сфере. Среди задач ГЭВУ также координация деятельности Европейского альянса по ИИ, поддержание диалога с мировым

сообществом, анализ экспертных мнений и подготовка докладов для Еврокомиссии. Полномочия группы не имеют срока окончания.

Вместе с тем анализ подготовленных ГЭВУ документов свидетельствует, что целью её работы уже изначально было определение не сущности «этичного ИИ», а норм и правил разработки основанных на нем систем. Только при их соблюдении такие системы могут быть признаны «доверенными» (trustworthy), пользователь якобы может использовать их без риска (а на деле, вероятно, не понимая этого риска, доверяясь различным авторитетам), руководствуясь как этическим императивом упомянутым выше тезисом Айзека Азимова.

Обеспечить доверие к самим предлагаемым ГЭВУ решениям Еврокомиссия предполагает за счет следования выработанным ею же правилу «трех шагов»: сначала необходимо определить ключевые требования к «доверенному ИИ», затем обсудить их с представителями промышленности и пользователями, и напоследок – провести международное согласование развития ИИ исключительно в интересах человека. Свой подход к развитию «этичного ИИ» ЕК намерена вынести на глобальное обсуждение.

Следуя изложенным соображениям, ГЭВУ разработала и 26 июня 2019 г. представила на первой европейской Ассамблее Альянса ИИ два документа, которые могут рассматриваться как основа политики ЕС в сфере ИИ.

Первый – «Руководящие указания по вопросам этики для доверенного ИИ» (Ethics Guidelines for Trustworthy Artificial Intelligence)[x], в котором эксперты ГЭВУ представили свое видение этических принципов доверия искусственному интеллекту (ИИ), опубликован 8 апреля 2019 года.

К основным принципам, которым должно отвечать заслуживающий доверия ИИ, авторами указаний отнесены:

- Законность – соблюдение всех норм законодательства, как в отношении самих целей внедрения, так и источников получения данных;
- Этичность – соблюдение этических ценностей и требований;
- Надежность с технической и социальной точки зрения.

В документе также определены семь требований к системам ИИ, позволяющие признать их «доверенными»:

- Надзор и контроль со стороны человека. Целью создания систем ИИ ставится расширение возможностей человека, оставляя за ним право принимать самостоятельные решения, не вводить в заблуждение. В них должны быть включены надлежащие механизмы надзора, которые могут быть достигнуты с помощью подходов «человек в контуре», «человек над контуром» и «человек в управлении»;
- Надежность и безопасность. Системы ИИ должны быть устойчивыми и технически безопасными. На случай сбоя или системных ошибок они должны предусматривать компенсирующую их запасную «модель поведения»;
- Конфиденциальность управления данными. Помимо обеспечения конфиденциальности и защиты данных, в системе ИИ должны быть также обеспечено наличие адекватных механизмов управления данными с учетом их качества и целостности, и законный доступ к данным. Люди должны иметь полный контроль над своими личными данными, их использование против воли человека и самого человека должно быть исключено;
- Многообразие, недискриминационность и справедливость. ИИ должен учитывать все человеческие особенности, возможности и потребности, и быть доступным для всех. Необходимо избегать несправедливой предвзятости и дискриминации,

поскольку они могут иметь многочисленные негативные последствия, начиная с маргинализации уязвимых групп и кончая обострением противоречий в обществе;

- Социальное и экологическое благополучие. ИИ должны приносить пользу всем людям, включая будущие поколения и не наносить вред окружающей среде. ИИ должен способствовать укоренению позитивных социальных изменений и усиливать экологическую ответственность;
- Подотчетность человеку. Системы ИИ должны предусматривать наличие механизмов обеспечения ответственности и отчетности за результаты их работы. Аудируемость, которая позволяет оценивать алгоритмы, данные и процессы проектирования, играет, по оценке экспертов ГЭВУ, ключевую роль в искусственном интеллекте.
- Прозрачность или интерпретируемость результатов. Бизнес-модели данных, результатов работы и самой системы ИИ должны быть прозрачными. Этому должны способствовать механизмы отслеживания. Кроме того, системы ИИ и их решения следует разъяснять таким образом, чтобы они были адаптированы к запросам соответствующих заинтересованных сторон. Люди должны знать, что они взаимодействуют с системой ИИ, и должны быть проинформированы о возможностях и ограничениях системы.

Однако в опубликованном также 8 апреля 2019 г. и дополняющем Руководящие указания документе, «Определение ИИ: основные возможности и научные дисциплины» («A definition of Artificial Intelligence: main capabilities and scientific disciplines»), опираясь на то, что «термин ИИ содержит явную ссылку на понятие интеллекта, а последнее (как в машинах, так и у людей) является расплывчатым понятием», ГЭВУ заявила, что в дальнейшем будет использовать только термин «система ИИ».

Такое уточнение является существенным, поскольку не только вносит терминологическую ясность, но делает логичными представленные в Руководящих указаниях требования к доверенному (этичному) ИИ. Последние становятся понятными, если относить их, как сказано в Определении к «системам, которые демонстрируют интеллектуальное поведение». Мало того, это подтверждает практическую направленность всей работы, но оставляет открытым вопрос, что же такое ИИ.

Чтобы упростить и упорядочить процесс реализации Руководящих указаний, ГЭВУ подготовила «Рекомендации по политике и инвестициям для доверенных ИИ» (Policy and investment recommendations for trustworthy Artificial Intelligence)[xi], одобренный 26 июня 2019 года. В нем содержатся 33 детализированные рекомендации, затрагивающие стимулирование, регулирование и финансирование развития и применения в ЕС технологий ИИ[xii].

Несмотря на название, основной упор в документе делается на социальные и этические вопросы. Эксперты ГЭВУ рекомендуют правительствам стран ЕС взять на себя обязательство поощрять разработку и применение систем ИИ, которые могут быть отнесены к категории доверенных, не использовать ИИ для контроля за поведением граждан и над их финансовым состоянием.

Рекомендации призывают правительства стимулировать внедрение «доверенных» систем ИИ, действующих под контролем человека в соответствии с нормами права и надёжны в прямом, т.е. техническом, смысле, а также предполагающих, в частности, защиту персональных данных, контроль за ними со стороны людей, которым эти данные принадлежат. Системы ИИ при взаимодействии с людьми должны идентифицировать себя во избежание ситуаций, когда, например, робот при онлайн-общении выдаёт себя за человека.

На самом деле и эти рекомендации, и 3 принципа с 7 требованиями из Руководящих указаний не являются какими-то откровениями и прямо следуют из принятого Первым симпозиумом по роботехнике в Сан-Ремо 2004 г. «Манифеста роботехники», а идеологически следуют из сформулированных еще Айзеком Азимовым законов.

Решением Ассамблеи Альянса ИИ с 1 июля до 1 августа 2019 г. запущена *пилотная* фаза тестирования закреплённых в Руководящих указаниях положений, в том числе в компаниях и организациях, связанных с разработкой и использованием систем ИИ. Её основная цель – сбор комментариев и предложений сообщества специалистов, в том числе из компаний и организаций, связанных с разработкой и использованием ИИ, в которых планируется опробовать разработанные ГЭВУ идеи, и доработка на их основе текста руководства.

Обратная связь обеспечена через открытый опрос или *количественный анализ*, материалы которого будут разосланы всем зарегистрировавшимся в пилотном проекте, или взаимодействие с рядом представительных организаций, чтобы получить более глубокие отзывы и комментарии по отдельным фрагментам руководства. Кроме того, Европейский Альянс ИИ через свои интернет-ресурсы предоставил постоянную возможность делиться с авторами отзывами и лучшими практиками.

19 февраля 2020 года Еврокомиссия представила Белую книгу по искусственному интеллекту - европейский подход к стандартам и ответственности, долгожданный программный документ, в котором излагаются конкретные меры по содействию развитию, внедрению и использованию ИИ, а также решению проблем, касающихся защиты фундаментальных прав[xiii].

С целью финансовой поддержки продвижения идеи мирового лидерства Европы в области ИИ за счет создания якобы этического ИИ Еврокомиссия на 70% увеличила ежегодные инвестиции в разработки ИИ, уже составляющие 1,5 млрд. евро в год. Европейский комиссар по вопросам цифровых технологий Мария Габриэль заявила однако, что если ЕС планирует стать пионером в области ИИ, этих 1,5 млрд евро хватит лишь на первоначальный этап, и в 2020 году на развитие ИИ Евросоюзу потребуется около 20 млрд евро. Решение нашлось с учетом аналогичных усилий государств - членов ЕС и частного сектора, на которые рассчитывают в Брюсселе, ориентиром рассматривается достижение общей суммы более чем в 20 млрд. евро в год в течение следующего десятилетия[xiv].

Без военных денег все равно не обойтись

Упорное обращение к этической стороне ИИ как товара, рекламируемого для вынесения на рынок, многими рассматривается не более как маркетинговый ход, определяемый тем, что, как заметил бывший вице-президент ЕК по единому цифровому рынку Андрус Ансип, «этическое измерение искусственного интеллекта не является предметом роскоши или чем-то таким, чем можно пренебречь ... этический ИИ - это выгодный всем проект, который может стать конкурентным преимуществом для Европы в разработке искусственного интеллекта, который создается для человека и которому люди могут доверять».[xv] Но, если вопрос о том, существует ли рынок ИИ, для многих пока открыт, то потребность в таком товаре на рынке уже сформирована. И большую помощь в этом оказали военные.

Американское издание *Defense News* опубликовало небезынтересное эссе заместителя министра обороны США по исследованиям и разработкам Майкла Гриффина (Michael D.Griffin) под заголовком «Вызов нашего времени» («The challenge of our time»), дающий наглядное представление об установках формирования технологической стратегии Соединенных Штатов, коррелированных в том числе с интересами Пентагона. «Некоторые потенциально дающие качественно новые возможности технологии, – отмечает Гриффин, – такие как искусственный интеллект..., движимы ныне не только оборонными заказами, а

коммерческими интересами и не только нашей отечественной [США – Авт.] промышленной базы»[xvi]. Это ли не стимул поддержать развитие ИИ?

Стремления Пентагона понятны и естественны. Однако в начале ноября 2019 Конгресс США обязал упоминавшуюся выше Комиссию национальной безопасности по искусственному интеллекту[xvii] вести постоянный анализ направлений развития ИИ с целью обеспечения требуемого уровня конкурентоспособности на мировой арене, достижения технологического превосходства Америки, укрепления сотрудничества с союзниками и партнерами, определения объемов необходимых ассигнований для проведения фундаментальных и прикладных исследований и поисковых разработок, постоянного анализа потенциальных рисков использования ИИ в военной сфере и определения перспективных направлений работ в этой области. При этом была принята позиция комиссии, ставящая главным приоритетом работ установление мирового лидерства США в области ИИ[xviii].

Не безразличны к исследованиям в области *этичного* ИИ и в европейских военных и оборонных ведомствах. Так, в министерстве вооруженных сил Франции, действуя, видимо, по правилу «и невинность соблюсти, и капитал приобрести», в поисках «золотой середины» «решили развивать оборонный искусственный интеллект в соответствии с тремя основными принципами: соблюдение норм международного права, поддержание достаточного контроля со стороны человека и обеспечение постоянной ответственности командования. Чтобы обеспечить ежедневное соблюдение этих принципов в долгосрочной перспективе и подкреплять нашу [Франции – Авт.] этическую мысль, так как новые виды использования ИИ появляются каждый день, [решили] создать министерский комитет по этике, занимающийся вопросами обороны. Этот комитет вступит в должность в самом конце [2019] года и окажет помощь в принятии решений и ожиданиях. Его основная роль будет заключаться в решении вопросов, возникающих в связи с новыми технологиями и их потенциальным использованием в оборонной сфере»[xix]. Вот так.

Но что-то же все-такистораживает...

Между тем в европейской практике естьстораживающие моменты, заставляющие серьезно задуматься над реальными целями проводимых работ.

На официальном сайте Еврокомиссии прямо указывается, что ИИ стал областью стратегического значения[xx]. Это утверждение следует признать верным, но при том понимании, что сфера ИИ затрагивает и вопросы безопасности. Отчасти (со стороны военных) мы это уже показали. Но, если военные разработки в основном ведутся вне Еврокомиссии на национальном уровне и в рамках НАТО, то в работах по применению ИИ в целях борьбы с преступностью ЕК тесно сотрудничает с комиссиями Совета Европы. Так, созданный в рамках СЕ специальный комитет по искусственному интеллекту нацелен на совершенствование правовой базы применения ИИ в интересах поддержки «Конвенции Совета Европы о киберпреступности»[xxi]. В настоящее время начата работа над дополнительным протоколом, регулирующим процедуру получения в ходе следственных действий в рамках самой спорной статьи Конвенции 32в) доказательств с серверов, находящихся в юрисдикции иностранного государства. Очевидным политико-правовым результатом этой работы станет легитимация тайного проникновения в киберпространство других государств с применением специальных (интеллектуальных) средств, против чего уже почти 20 лет выступают Россия, Китай и ряд других стран.

На внеэкономическую направленность работ над *этичным* ИИ указывает и циркулирующее в среде экспертов мнение о том, что идеи лидерства Европы на мировом рынке не более чем заблуждение, основанное, как минимум, на трех других: что существует рынок ИИ, что другие страны не заинтересованы в моральности ИИ, и что европейцы

имеют конкурентное преимущество в производстве систем ИИ, которые являются более этичными, чем произведенные в другом месте. Такую позицию занимает, в частности, Директор упомянутого выше Центра инноваций данных ЕС и вице-президент Фонда информационных технологий и инноваций Даниэль Кастро, считающий, что разработка этичных систем ИИ важна, но идея доминирования ЕС на мировом рынке ИИ, когда он будет, исключительно путем разработки наиболее этичных систем, не подтверждается фактами. По его мнению, ИИ – это обобщение, подразумевающее лишь следующую волну инноваций в системах управления всех уровней[xxii].

Развивая свою мысль, Кастро обращает внимание на то, что ЕС не одинока в рассмотрении этичности ИИ. У многих стран есть политика решения этой проблемы. В США, например, разработаны технические стандарты для ИИ, в том числе для поддержки его этического использования. Более того, госструктуры США, такие как министерство обороны и министерство транспорта, выступили с собственными инициативами по обеспечению этического использования ИИ в своих соответствующих областях. В Китае поддерживаемая правительством Китайская ассоциация искусственного интеллекта разработала свои собственные этические руководящие принципы.

Основания для скепсиса дают и результаты научных исследований по робоэтике. В принятом еще в 2004 г. на Первом Международном симпозиуме по робоэтике[xxiii] в Сан-Ремо «Манифесте робоэтики» отмечено, что робоэтика не этика роботов, а соблюдение этических норм человеком при разработке, создании и использовании робототехнических систем. Это, безусловно, распространяется и на «этичный» ИИ, а разработанные экспертами ГЭВУ в Руководящих принципах семь требований к доверенным системам ИИ суть требования не к системам, а к их разработчикам.

По разным оценкам, до 40% компаний, использующих термин ИИ в описании своих проектов, делают это только для привлечения инвесторов. Сами предприниматели отмечают непонимание ими сущности ИИ и тем более его моральности. В среде аналитиков появилась оценка, что европейский «сектор ИИ — это пузырь, который скоро может лопнуть». «И пока неизвестно, – отметил в уже упоминавшемся выступлении на 74 сессии ГА ООН Борис Джонсон – станет ли искусственный интеллект благом для человечества или породит «терминаторов с красными глазами», готовых его уничтожить».

Дополнительную настороженность в восприятии европейских подходов к разработке этичного ИИ вызывает и то, что, приняв новое назначение, Урсула фон дер Ляйен призвала «сделать все для объединения усилий Евросоюза и США, чтобы вместе противостоять России, а не бороться с ней отдельно друг от друга», и продолжить придерживаться санкций в отношении Москвы «с позиции силы»[xxiv].

Кроме того, уже звучат призывы не допускать «недемократические страны» к обладанию и использованию технологий ИИ. В настоящее время комиссии Парламентской Ассамблеи Совета Европы (ПАСЕ) работают над подготовкой пяти докладов о необходимости создания демократических механизмов контроля над использованием ИИ. Обсуждение их результатов в ПАСЕ запланировано на 2020 год.

Интересно отношение к проблеме Римской католической церкви. На прошедшей в середине октября 2019 г. в Ватикане под эгидой Папы Франциска и с его участием прошла конференция по проблемам ИИ. Среди приглашенных были представители компаний из Силиконовой долины, в частности Facebook, Mozilla и Western Digital, государственных органов и инвестиционных банков, специалисты по компьютерным технологиям и католической этике. Высказанная на ней позиция Понтифика не допускает разночтений:

«Впечатляющие разработки в области технологий, особенно в сфере искусственного интеллекта, начинают оказывать все более существенное влияние на все сферы человеческой деятельности. Поэтому открытые и конкретные дискуссии на эту тему необходимы как никогда». Папа заявил, что дабы они не могли быть использованы в злонамеренных целях, технологии нуждаются в теоретических и практических моральных принципах. ИИ сможет «отравить» тему публичных дебатов, распространять ложные мнения и манипулировать миллионами людей одновременно. «Может даже существовать угроза учреждениям, которые должны гарантировать мирное сосуществование», – предупредил он[xxv]. В Ватикане прошедшая встреча рассматривается более чем серьезно – материалы конференции будут использованы для подготовки специальной папской энциклики о технологиях искусственного интеллекта.

И все-таки война?

Немаловажным моментом является также то, что весь этот активно выносимый в СМИ и публичные дебаты дискурс об этических, человеко-ориентированных ИИ-системах как бы оставляет в тени их прямое военное применение и соответствующие разработки[xxvi], проводимые в военных исследовательских центрах. Между тем только Министерство обороны США планирует в 2020 году потратить на исследования в области ИИ 4,5 млрд. дол.[xxvii] По оценкам Пентагона, в 2020 году вложения Пекина в свой военный ИИ-потенциал достигнут 60 млрд. долларов[xxviii].

В свою очередь, ЕК, стараясь не отставать от лидеров, приняла решение о создании Европейского оборонного фонда. Хотя в полной мере фонд начнет функционировать с 2021 года, через него уже в 2019 г. на поддержку только совместных оборонно-промышленных проектов, финансируемых через Европейскую программу оборонно-промышленного развития, выделено 25 млн. евро в дополнение к 500 млн., изначально выделенным на 2019-2020 годы. При этом в официальном сообщении ЕК поясняется, что в финансируемых проектах «основное внимание будет уделяться таким областям, как ... искусственный интеллект»[xxix]. Этическую суть такого решения выразила Флоранс Парли министр вооруженных сил Франции: «ИИ не является самоцелью; он должен способствовать более информированному и более быстрому принятию решений в интересах наших солдат... ИИ также означает лучшую защиту наших войск... Наконец, ИИ означает более сильную киберзащиту. Кибер-солдаты смогут с очень высокой скоростью противостоять все более скрытным, многочисленным и автоматизированным атакам, которые угрожают нашей экономике»[xxx]. Надо полагать, любой военный министр стран ЕС присоединился бы к своему французскому коллеге, если бы его об этом спросили. Но почему-то участники упомянутой конференции в Ватикане выразили серьезную озабоченность тем, что дальнейшие разработки в этой сфере могут спровоцировать новую гонку вооружений. «Машине нельзя доверять решение, будет ли человек жить или умрет»[xxxi].

Так и хочется сказать: «Читайте классиков». Технологии меняются, а политика остается.

[i] Краткий, хотя и несколько устаревший (октябрь 2018 г.) обзор читатель может найти в статье Dutton T. An Overview of National AI Strategies [Electronic resource] / Medium. 2018. Mode of access: <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>

[ii] Griffin M. Pentagon R&D boss: The challenge of our time [Electronic resource] / Defense News. 2019. Mode of access: <https://www.defensenews.com/outlook/2019/12/02/pentagon-rd-boss-the-challenge-of-our-time/>

[iii] По оценкам европейских аналитических центров, в частности Центра инноваций данных, ЕС отстает в сфере ИИ не только от США, но и от Китая и России. В ЕК обращают внимание на высказанное еще в 2013 г. Бараком Обамой предположение, что работы по ИИ будут не менее эффективными, чем вложения в геном человека, где 1 доллар инвестиций принес в экономику США 140 долларов (См. Попсулин С. США визуализируют работу человеческого мозга [Электронный ресурс] / CNews. 2013. URL: https://www.cnews.ru/news/top/ssha_vizualiziruyut_rabotu_chelovecheskogo (дата обращения: 09.04.2020).

[iv] RB New Commission President Promises Legislation for Ethical AI [Electronic resource] / Medium. 2019. Mode of access: <https://medium.com/@fourthquadrant/new-commission-president-promises-legislation-for-ethical-ai-3fa73cfc5d43>

[v] Из выступления Г.Грефа на Международной конференции по искусственному интеллекту (Artificial Intelligence Journey) 09.11.2019. См. Колесников А. Заслуженные деятели искусственного: как Владимир Путин оцифровал Германа Грефа // Газета "Коммерсантъ" №206 от 11.11.2019, стр. 1. URL: <https://www.kommersant.ru/doc/4154715> (дата обращения: 09.04.2020).

[vi] EU Member States sign up to cooperate on Artificial Intelligence [Electronic resource] / European Commission. 2018. Mode of access: <https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence>

[vii] Open letter Research priorities for robust and beneficial artificial intelligence [Electronic resource] / Future of Life Institute. Mode of access: futureoflife.org/ai-open-letter

[viii] von der Leyen U. A Union that strives for more. My agenda for Europe / Political Guidelines. Mode of access: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

[ix] Здесь Европа не была оригинальна: 15 мая того же года решением S.2806 Сенат США учредил Комиссию национальной безопасности по искусственному интеллекту (National Security Commission on Artificial Intelligence). Ее задачей стал обзор достижений в области ИИ, машинного обучения и связанных с ними технологий. Руководство КНБ ИИ возложено на бывшего генерального директора Google Эрика Шмидта.

[x] Ethics guidelines for trustworthy AI [Electronic resource] / European Commission. 2019. Mode of access: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

[xi] Policy and investment recommendations for trustworthy Artificial Intelligence [Electronic resource] / European Commission. Mode of access: <https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>

[xii] Идеологически эти рекомендации, как и содержащиеся в Руководящих указаниях, не являются принципиально новыми и восходят к Манифесту роботетики, принятому на Первом симпозиуме по роботетике (Сан-Ремо 2004 г.)

[xiii] White Paper on Artificial Intelligence – A European Approach to Excellence and Trust / European Commission. Mode of access: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[xiv] Artificial Intelligence [Electronic resource] / European Commission. Mode of access: <https://ec.europa.eu/digital-single-market/en/artificial-intelligence>

[xv] Artificial intelligence: Commission takes forward its work on ethics guidelines [Electronic resource] / European Commission. Mode of access: https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1893

[xvi] Griffin M. Pentagon R&D boss: The challenge of our time [Electronic resource] / Defense News. 2019. Mode of access: <https://www.defensenews.com/outlook/2019/12/02/pentagon-rd-boss-the-challenge-of-our-time/>

[xvii] National Security Commission on Artificial Intelligence. Создана решением Сената США S.2806 от 15 мая 2018 г. на период до 1.10.2020 как независимая комиссия для обзора достижений в области ИИ, машинного обучения и связанных с ними технологий.

[xviii] Владимир Иванов. Искусственный интеллект – основа будущих сражений // Независимое военное обозрение №42(1070). 2019. URL: http://nvo.ng.ru/concepts/2019-11-15/1_1070_al.html (дата обращения: 09.04.2020).

[xix] Parly F. France's armed forces minister: How AI figures into operational superiority [Electronic resource] / Defense News. 2019. Mode of access: <https://www.defensenews.com/outlook/2019/12/02/frances-armed-forces-minister-how-ai-figures-into-operational-superiority/>

[xx] Artificial Intelligence [Electronic resource] / European Commission. Mode of access: <https://ec.europa.eu/digital-single-market/en/artificial-intelligence>

[xxi] Россия не является ее участником и в последние годы продвигает в рамках ООН новый альтернативный вариант Конвенции

[xxii] Castro D. Europe will be left behind if it focuses on ethics and not keeping pace in AI development [Electronic resource] / Euronews. 2019. Mode of access: <https://www.euronews.com/2019/08/07/europe-will-be-left-behind-if-it-focuses-on-ethics-and-not-keeping-pace-in-ai-development>

[xxiii] [First International Symposium on Roboethics](http://www.roboethics.org/sanremo2004/). January 30-21, 2004. Mode of access: <http://www.roboethics.org/sanremo2004/>

[xxiv] von der Leyen U. A Union that strives for more. My agenda for Europe / Political Guidelines. Mode of access: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf

[xxv] Папа Римский предупредил об опасностях искусственного интеллекта // Вести-RU. 2019.

URL: <https://www.vestifinance.ru/articles/125891> (дата обращения 09.04.2020)

[xxvi] Мы оставили за скобками нашего рассмотрения другую на сегодня важнейшую и, безусловно связанную с ИИ проблему ограничения военного применения ИИ – «смертоносных автономных систем», более известных как «роботы-убийцы». Этот вопрос, как представляется, достаточно системно рассматривается в специальной литературе, и военными, и политологами. Кроме того, по этой теме уже несколько лет ведется международная дискуссия в рамках переговорного процесса по «Конвенции о негуманном оружии» (КНО), создана и функционирует Группа правительственных экспертов и т.д. В общем, можно констатировать, что «процесс пошел». Чем дело кончится, сказать трудно – проблема сложная и не проработанная даже теоретически. Интересующимся можно было бы порекомендовать, в частности, работы сотрудников ПИР-Центра В.Козюлина [Боевые роботы: \(смертоносные автономные системы\) – военные, политические, правовые, этические вызовы; Смертоносные автономные системы вооружений: проблемы современного международно-правового регулирования и перспективы их решения; Три](#)

[группы угроз смертоносных автономных систем](#) и О.Демидова ([Как ограничить смертоносные автономные системы](#), НИУ ВШЭ К. Мартынова [Этика автономных машин: деонтология и военные роботы](#)), а также материалы конференций по КНО.

[xxvii] В выше цитировавшейся статье Майка Гриффина замглавы Пентагона, сетуя на потерю США лидерства в ключевых технологиях, прямо говорит: «Мы знаем, что нам нужно делать. В Стратегии национальной обороны указаны основные направления инвестиций, которые мы должны сделать: возрождение ядерной триады, микроэлектроника, кибербезопасность, биотехнология, 5G, космос, гиперзвуковые технологии, искусственный интеллект, направленная энергия, автономные системы, сетевые коммуникации, противоракетная оборона, квантовая наука и другие. Превосходство в этих технологиях, вплетенных в новую военную архитектуру, которая бросает вызов нашим противникам, а не реагирует на их действия.