

УДК 327.7

## **ДИАЛЕКТИКА ЭТИКИ И ПРАВА В РЕГУЛИРОВАНИИ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ОПЫТ ЕС**

© 2022 **МАСЛОВА Елена Александровна\***

*Кандидат политических наук, доцент*

*Кафедра интеграционных процессов, ИМИ МГИМО МИД России*

*119454, г. Москва, проспект Вернадского, 76*

*Отдел Черноморско-Средиземноморских исследований, Институт Европы РАН*

*125009, г. Москва, Моховая ул., дом 11, стр. 3*

**\*E-mail:** *e.maslova@inno.mgimo.ru*

© 2022 **СОРОКОВА Екатерина Дмитриевна\*\***

*Аспирант МГИМО МИД России*

*119454, г. Москва, проспект Вернадского, 76*

**\*\*E-mail:** *sorokova.e.d@my.mgimo.ru*

Поступила в редакцию 19.02.2022

После доработки 26.05.2022

Принята к публикации 08.07.2022

**Аннотация.** Европейский союз намерен стать лидером по разработке систем искусственного интеллекта (ИИ). Внедрение технологий ИИ обладает трансформирующим потенциалом и представляет собой вызов – не только в глобальном смысле, но и для евробюрократии, перед которой встает дилемма регулирования ИИ. Суть этой дилеммы состоит в том, чтобы, не сдерживая технический прогресс, построить модель управления, минимизирующую неопределенность и риски катастрофических последствий для государства, общества и человека. В статье представлены этический и правовой подходы ЕС к регулированию ИИ. Проанализированы основные этические принципы ЕС в сфере ИИ и их эволюция; представлен анализ нормативно-правового регулирования ИИ через теорию сложных систем и многоуровневого управления. Сделан вывод, что сегодня ЕС отдает приоритет «доверенному ИИ», однако в последнее время наблюдается постепенный переход от «мягких» форм регулирования к более обязывающим. При этом руководящая роль в этих процессах сохраняется за Еврокомиссией.

**Ключевые слова:** этика международных отношений, этика искусственного интеллекта, искусственный интеллект, ИИ, Европейский союз, сложные системы.

**DOI:** 10.31857/S0201708322050023

**EDN:** glofsa

Профессор Гарвардского университета С. Хоффманн писал, что наличие правил преобразует международную среду «из состояния джунглей в состояние международного общества» [Hoffmann, 1981]. Однако «джунглификация» международных отношений в той или иной степени неизбежна – не только из-за нарушения уже принятых норм, но и из-за изменения самой среды. Развитие новых технологий, в частности, технологии искусственного интеллекта (ИИ) ставит мировое сообщество перед новым вызовом. Возникают озабоченности в сфере безопасности, прежде всего, военной и инфраструктурной, связанной с недружественным применением ИИ и его непредсказуемыми последствиями. Одновременно происходит рост внимания к социально-гуманитарным проблемам, связанным с воздействием ИИ на индивида и общество. Уже запущенная цифровая трансформация на основе ИИ начинает восприниматься как глобальный вызов. Это подтверждают ежегодные доклады Всемирного экономического форума, где начиная с 2015 г. отдельное внимание уделяется рискам, связанным с новейшими технологиями и ИИ<sup>1</sup>.

Основным вопросом и задачей для мирового сообщества становится регулирование применения ИИ в различных сферах, где инструментами выступают правовые и моральные нормы. Несмотря на всю неоднозначность регулирующей роли права и морали в международных отношениях и мировой политике, мир не знает других инструментов [Цыганков, 2003: 349]. Так, с одной стороны, происходит усиление роли этических факторов при возрастающем уровне правового сознания, с другой – их несоблюдение («джунглификация»).

Диалектика права и этики состоит в том, что оба представляют *правила поведения*, «приобретающие роль юридического или нравственного долга и ответственности за его нарушение, отражающие существующий уровень развития международной системы, человеческой цивилизации в целом» [Цыганков, 2003: 373]. При этом нормы формируют не просто принципы поведения, но и конструируют предпочтения, идеалы, идентичность (конститутивная функция [Истомин, 2021: 88]). В конечном итоге распространение общих норм способствует укреплению стабильности международной системы (во многом за счет предсказуемости).

### ***Поиск модели регулирования ИИ: к постановке проблемы***

Вопрос законодательного регулирования в сфере ИИ крайне актуален для стран, активно занимающихся разработками этих технологий. В условиях нормативного

---

<sup>1</sup> Global Risks (2015) 10th Edition. World Economic Forum. URL: [https://www3.weforum.org/docs/WEF\\_Global\\_Risks\\_2015\\_Report15.pdf](https://www3.weforum.org/docs/WEF_Global_Risks_2015_Report15.pdf) (дата обращения: 17.05.2022)

вакуума интенсивное качественное развитие, внедрение во все сферы жизни общества и реализация трансформирующего потенциала ИИ могут радикально усложнить ландшафт технологических рисков современности.

Еще ни в одной стране ЕС не принят всеобъемлющий закон, охватывающий общие принципы проектирования, производства и эксплуатации систем ИИ. В большинстве случаев на технологии ИИ распространяются положения нормативных актов из смежных сфер (например, правила обработки и защиты данных; прав человека/потребителя/на частную жизнь; правила разработки, тестирования и использования автономных транспортных систем). В странах, активно развивающих ИИ (например, во Франции, Германии), параллельно действуют инструменты «мягкого» регулирования – кодексы этики, кодексы поведения, руководства, «белые книги» и иные документы, содержащие руководящие принципы и рекомендации для разработчиков систем ИИ. Помимо государства, крупные мировые компании в сфере ИИ также создают собственные принципы ответственности (*Microsoft, Google AI*). По данным исследования из Обсерватории ИИ ОЭСР, в 2006–2019 гг. на территории Европы появилось 70 документов «мягкого права», касающихся ИИ, причем в подсчете учтены только документы, выпущенные институтами ЕС, Советом Европы, НКО и другими акторами, без учета инициатив отдельных стран. Наиболее популярны такие форматы документов, как рекомендации и стратегии (более 54%), принципы (почти 25%) и стандарты (9%) [Gutierrez and Marchant, 2021].

В условиях стремительного развития самой технологии *in se* задача законодателей становится нетривиальной: *не сдерживая технический прогресс, построить модель управления, минимизирующую неопределенность и риски катастрофических последствий для государства, общества и человека*. Она дополнительно осложняется тем, что предложения о введении законодательного и даже этического регулирования встречают сопротивление со стороны разработчиков и некоторых экспертов, которые видят в этом ограничения для развития.

Эффективная модель регулирования ИИ и взаимодействия различных акторов должна учитывать нелинейные процессы развития ИИ-технологий и связанных с их внедрением изменений в экономической, социальной и политической сферах. Наиболее полно учесть все эти факторы при анализе как самих процессов, так и политики их регулирования, позволяет современная *теория сложных систем* [Истомин, Дегтярев, 2015; Cairney, 2012].

ЕС намерен стать лидером по разработке систем ИИ, что является одним из заявленных направлений деятельности ЕС в развитии ИИ-технологий. Европейский подход направлен на обеспечение того, чтобы развитие ИИ было основано на «правилах, предоставляющих гарантии безопасности для функционирования рынков и государственного сектора, безопасность и основные права людей»<sup>1</sup>. Таким образом, ЕС сталкивается с вызовом регулирования технологии ИИ, где в основании подхо-

<sup>1</sup> A European approach to artificial intelligence (2022). European Commission. URL: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> (дата обращения: 17.05.2022)

да лежит вопрос взаимодополняемости этического и правового подходов. Целью статьи является комплексный анализ регулирования сложных систем ИИ ЕС через призму этих двух подходов.

В отечественной научной литературе вопросы регулирования ИИ рассматривали такие эксперты, как А.В. Незнамов [Незнамов, 2019], М.В. Федоров, Ю.Д. Цветков [Федоров, Цветков, 2020], А.В. Абрамова, А.А. Кулешов [Kuleshov A. et al., 2020]. В аналитическом докладе «Этика и цифра: от проблем к решениям» международному и национальному этическому регулированию ИИ посвящена отдельная глава [Доверенный..., 2021]. Политика ЕС в области ИИ, неоднозначное отношение Брюсселя к феномену ИИ и дискуссии о степени его регулирования стали темой исследования О.Ю. Потемкиной [Потемкина, 2019]. Кроме того, тема вызовов и рисков, требующих регулирования ИИ, освещается в связи с проблематикой безопасности. Например, Д.Ю. Базаркина исследует влияние ИИ на информационно-психологическую безопасность [Базаркина, Пашенцев, 2019], а также подходы ЕС к отражению гибридных угроз [Базаркина, 2021].

### ***Этические принципы ИИ: «кодифицировать нельзя зарегулировать»***

Разработка этических принципов ИИ изначально позиционировалась в стратегических документах ЕС как один из приоритетов. Так, в Сообщении Еврокомиссии (ЕК) «Об искусственном интеллекте для Европы» (2018) было обозначено первое официальное определение ИИ и три основных направления его развития: 1) экономическое (увеличение технологического и производственного потенциала ЕС и рост потребления ИИ в экономике); 2) социальное (подготовка к общественным и экономическим изменениям, вызываемыми ИИ); 3) нормативное (обеспечение этической и законодательной базы). Под ИИ подразумеваются системы, демонстрирующие разумное поведение, способные анализировать окружающую среду и совершать действия с некоторой долей автономии для достижения определенных целей<sup>1</sup>, а также программное обеспечение, способное к логическому рассуждению, накоплению знаний, разумному планированию, обучению, коммуникации, восприятию и манипулированию объектами<sup>2</sup>. Подразумевается, что ИИ может действовать как исключительно в виртуальном мире (голосовые помощники; поисковики; приложения с функцией распознавания лиц и речи), так и как часть устройства (сложная робототехника; автономно управляемые машины; дроны; «Интернет вещей»).

Нормативное направление предусматривало разработку Руководства по этике в сфере ИИ, в котором планировалось охватить несколько основных тем: будущее

<sup>1</sup> Communication from the Commission. Artificial Intelligence for Europe (2018). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN> (дата обращения: 17.05.2022) p. 1.

<sup>2</sup> Commission Staff Working Document ‘Liability for emerging digital technologies’ (2018). European Commission. URL: <https://ec.europa.eu/digital-single-market/en/news/european-commission-staff-working-document-liability-emerging-digital-technologies> (дата обращения: 17.05.2022) p. 23.

сферы труда; справедливость; безопасность; общественная интеграция; прозрачность; влияние ИИ на демократические принципы и применение Хартии основных прав человека, в т.ч. права на защиту частной жизни и достоинства, персональных данных, прав потребителей, а также принцип недискриминации. В Сообщении также отмечалось намерение ЕК следить за развитием регуляторики в сфере этики ИИ и при необходимости обновлять существующие правовые нормы для адаптации их к новым вызовам и обеспечения соблюдения базовых ценностей и фундаментальных прав ЕС<sup>1</sup>.

Реализация Европейской стратегии по ИИ была возложена на *AI HLEG* – Экспертную группу высокого уровня по ИИ (далее – ЭГВУ), состоящую из 52 экспертов – представителей компаний, научного сообщества и гражданского общества. В задачи Группы входила разработка рекомендаций для ЕК по работе с возможностями и вызовами ИИ в средне- и долгосрочной перспективе, которые легли бы в основу разработки политики, процесса оценки законодательства и создания цифровой стратегии нового поколения, а также создание Руководства по этике в сфере ИИ, проекта политической и инвестиционной стратегии по ИИ и работа с участниками Европейского Альянса по ИИ (площадки для диалога с гражданским обществом). Помимо Альянса, ЭГВУ работала над Рекомендациями, проходила совместно с Европейской группой по этике науки и новых технологий, подготовившей доклад по робототехнике и «автономным» системам (2018)<sup>2</sup>, и Группой высокого уровня в рамках инициативы по цифровизации европейской промышленности, на которую возложена координация работы с представителями стран-членов ЕС<sup>3</sup>.

Этическое руководство изначально позиционировалось не как список общих политических рекомендаций, а как набор практически применимых принципов, применимых в различных ситуациях и затрагивающих основные сферы, в которых необходимо определить этические ориентиры или провести дополнительные исследования. Предполагалось, что Руководство станет документом «мягкого права», и организации будут добровольно брать на себя обязательства по соблюдению правил, прописанных в Руководящих принципах.

Окончательная версия Руководящих принципов [Ethics Guidelines..., 2019] по этике для надежного («доверенного») ИИ была опубликована в апреле 2019 г. и состояла из трех частей. В первой обозначены основные этические принципы и ценности, которые необходимо соблюдать при развитии, использовании и внедрении технологий ИИ: уважение автономии человека, предотвращение вреда, справедливость и объяснимость. Подчеркивается необходимость особого внимания к уязви-

<sup>1</sup> Communication from the Commission. Artificial Intelligence for Europe (2018). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN> (дата обращения: 17.05.2022) p. 1.

<sup>2</sup> Ethics Guidelines for Trustworthy AI (2019). High-Level Expert Group on Artificial Intelligence. 40 p. URL: <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1> (дата обращения: 17.05.2022)

<sup>3</sup> Concept Note. The High-Level Expert Group on Artificial Intelligence (2018). URL: <https://futurium.ec.europa.eu/en/european-ai-alliance/open-library/concept-note-ai-hleg?language=en> (дата обращения: 17.05.2022)

мым группам и ситуациям, в которых возникает асимметрия властных полномочий (между работодателями и сотрудниками, предпринимателями и потребителями). Присоединяющиеся к Рекомендациям акторы признают, что помимо значимой пользы обществу и человеку, системы ИИ могут представлять риск и оказывать негативное воздействие, в т.ч. на демократию, верховенство права, правосудие и собственно человеческий интеллект, – в связи с этим принимают соответствующие меры для снижения риска.

Во второй части приведены *семь требований к «доверенному» ИИ*: 1) акторность и контроль человека; 2) техническая надежность и безопасность; 3) защита частной жизни и управление данными; 4) прозрачность; 5) разнообразие, недискриминация и справедливость; 6) благосостояние окружающей среды и общества; 7) подотчетность.

В третьей части Рекомендаций предпринята попытка операционализировать этические требования и подчеркнуть необходимость дальнейшей работы над их улучшением и адаптацией к различным системам ИИ. Работа по методологическому обеспечению реализации этических принципов продолжилась: еще год заняло создание методики оценки надежности ИИ (*ALTAI*), ее тестирование и публикация доработанного варианта.

В апреле 2020 г. ЕК выпустила Белую книгу [White Paper..., 2020], где, среди прочего, были обозначены направления совершенствования законодательства в сфере ИИ и представлен *риск-ориентированный подход*, в котором степень риска предлагалось оценивать, ориентируясь и на сектор, и на предполагаемое использование самой системы. В ходе общественных консультаций представленный подход подвергся серьезной критике как со стороны НКО, указавших на недостаточное внимание к проблеме защиты прав человека и узость секторального подхода, так и со стороны 14 государств, опасавшихся отнесения чрезмерного количества ИИ-технологий к классу высокорисковых<sup>1</sup>.

В 2021 г. ЕК представила новый пакет документов: Сообщение о поддержке европейского подхода к ИИ, Обновленный скоординированный со странами-участницами ЕС план по развитию ИИ и законопроект – Регламент о гармонизации правил, применяемых к системам ИИ. Основные приоритеты и направления деятельности остались неизменными: создание благоприятных условий для внедрения и развития ИИ; развитие передовых технологий ИИ в ЕС от лабораторий до выхода на рынок; обеспечение того, чтобы ИИ стал силой, работающей на благо людей и общества; а также формирование стратегического лидерства в секторах с высокой отдачей<sup>2</sup>.

<sup>1</sup> Europe's Approach to Artificial Intelligence: How AI Strategy is Evolving. Access Now, 2020. URL: <https://www.accessnow.org/cms/assets/uploads/2020/12/Europes-approach-to-AI-strategy-is-evolving.pdf> (дата обращения: 17.05.2022), p. 24.

<sup>2</sup> A European approach to artificial intelligence (2022). European Commission. URL: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence> (дата обращения: 17.05.2022)

Таким образом, вектор развития ИИ в ЕС остался принципиально неизменным, но именно в 2021 г. регулирование ИИ в ЕС вышло на новый уровень. Примечательно, что роль мер саморегулирования учтена при разработке нормативно-правового регулирования, создание которого стало закономерным продолжением политики по установлению «правил игры».

### ***Регламент о гармонизации правил, применяемых к системам ИИ, через призму теории сложных систем***

В апреле 2021 г. ЕК представила первый комплексный законопроект – Регламент о гармонизации правил, применяемых к системам ИИ [Proposal for a Regulation..., 2021] (далее – Регламент), предполагающий применение более жестких форм контроля к высокорисковым системам ИИ и более гибких инструментов регулирования для систем, не связанных с высоким риском.

Профессор Абердинского университета П. Кэйрни, применивший теорию сложности для анализа публичной политики и процессов управления, выделяет шесть основных характеристик сложных систем [Cairney, 2012]. Опираясь на выделенные им основные характеристики сложных систем, рассмотрим предлагаемую в законопроекте риск-ориентированную модель регулирования и систему взаимодействия различных акторов.

1) Основное внимание при изучении сложных систем уделено *динамике процессов и взаимодействий как внутри системы, так и между системой и внешней средой*. При этом анализ системы не сводится к анализу свойств ее элементов. Данное положение применимо и к системе регулирования ИИ, где на процесс оказывает влияние множество сил, а акторы нередко выступают в различных ролях, вынуждены быстро адаптироваться к постоянно меняющимся условиям и опираться не на единственную стратегию. Например, государство занимается не только нормотворчеством, но и контролем, одновременно выступает как инвестор, двигатель цифровизации, пользователь ИИ-технологий и т.д.

В случае ЕС основную часть этих функций берет на себя ЕК, но значительные полномочия по имплементации, уточнению и контролю возлагаются на государства-члены ЕС. Положения Регламента распространяются и на институты, агентства и иные структуры ЕС, действующие в качестве провайдеров или пользователей систем ИИ. При этом и объектом, и субъектом регулирования выступают разработчики ИИ-технологий, бизнес, индивиды, гражданское общество, НКО. Отдельно в контексте отношений системы с внешней средой отметим заложенный в Регламенте *принцип экстерриториальности*: системы ИИ подпадают под действие законопроекта, даже если они не имеют непосредственного доступа на рынок товаров и услуг ЕС и используются не на территории ЕС. То есть, например, если европейская компания заключает контракт с игроком из третьей страны, не входящей в ЕС, и оказываемые по договору услуги предполагают использование высокорисковых систем и каким-либо образом затрагивают граждан ЕС, то (в текущей редакции документа) на игрока из третьей страны также распространяются нормы Регламен-

та. Таким образом, в системе, изложенной в проекте Регламента, в той или иной степени затронуты все основные группы акторов ИИ.

2) *Нелинейный характер процессов* – повышает непредсказуемость последствий внутренних и внешних воздействий на систему и ее элементы. Масштабы причин и вызываемых ими следствий могут быть непропорциональны.

Данная характеристика также отражена в проекте Регламента. Примечательно, что законодатели и специалисты, занимающиеся проблематикой этичного использования ИИ, уже сейчас задумываются о масштабности возможных последствий и кумулятивного эффекта от неявного влияния систем ИИ на человека, и от множества нерегулируемых воздействий в особо чувствительных сферах. Высокорисковыми считаются системы, затрагивающие здоровье, безопасность и основные права человека. Среди них – используемые в критической инфраструктуре, образовании или профессиональном обучении, оказании частных и государственных услуг, при решении вопросов миграции, предоставления убежища и пр. Особенно вредоносными названы практики использования ИИ в целях манипуляций, эксплуатации и социального контроля.

В ст. 5 предлагается запретить, как противоречащие ценностям ЕС, системы ИИ, нацеленные на: а) искажение поведения человека, при котором возможно нанесение физического и психологического вреда (в т.ч. воздействия на подсознание, не воспринимаемые человеком или эксплуатирующие уязвимости детей или других людей в связи с их возрастом, физической или ментальной недостаточностью); б) создание социального скоринга<sup>1</sup>; в) биометрическую идентификацию людей в режиме реального времени (с определенными оговорками, связанными с правоохранительной и антитеррористической деятельностью).

3) *Восприимчивость к изначальным условиям*, что повышает вероятность возникновения зависимости от изначального пути развития.

Фактически это замечание П. Кейрни характеризует современный этап развития регулирования, предшествующий массовому распространению ИИ-технологий. Перед законодателями открывается возможность заранее выстроить и закрепить систему оценки и минимизации рисков, связанных с развитием и использованием ИИ. Дифференцированный подход к ИИ-технологиям – возможность задать вектор развития, обозначив приоритеты в сфере безопасности, и избежать, с одной стороны, ограничения технического прогресса, с другой – пагубных последствий бесконтрольного внедрения и расходов на ликвидацию таких последствий. Последнее актуально как для государства, ответственного за благополучие граждан, так и для компаний, которые не только понесут репутационные издержки, но и потратят средства на прохождение комплаенс-процедур.

Однако различные акторы рассматривают подобную зависимость от условий законодательного регулирования на ранних этапах развития ИИ не только как преимущество, но и как недостаток. В докладе «Сколько будет стоить Европе Акт об ИИ?» проект Регламента оценивается «как самый рестриктивный в мире инстру-

<sup>1</sup> Система социального рейтинга, предполагающая автоматизированную оценку действий граждан по определенным параметрам и формирование рейтинга.

мент регулирования ИИ», который «не только ограничит развитие и использование ИИ в Европе, но и возложит значительные расходы на бизнес и потребителей в Евросоюзе» [Mueller, 2021]. По расчетам авторов, вступление Регламента в силу к 2025 г. приведет к снижению инвестиций в сферу ИИ на 31 млрд евро, обернется повышением расходов для европейского бизнеса, снизит темпы цифровой трансформации и подорвет конкурентоспособность экономики ЕС. Похожие опасения в 2020 г. высказывали США и корпорация «Google», призывая ЕС отказаться от введения чрезмерного контроля над ИИ, поскольку такой путь в долгосрочной перспективе мог бы дать конкурентные преимущества странам с менее строгими законодательными режимами [Burns, Bradley, Powell, 2021].

4) *Самоорганизация, формирование в системе упорядоченных структур без целенаправленного воздействия со стороны внешних факторов или центра системы.*

Как было отмечено ранее, процесс создания этических норм в форме принципов, кодексов и иных документов мягкого права в европейском регионе активно начался в 2017 г. ЕК, в свою очередь, уловила тенденцию и придала ей ускорение, сделав одним из приоритетов цифровой политики ЕС. Таким образом, процесс формирования общеевропейского подхода начал идти параллельно с развитием саморегулирования, т.е. продолжающимися процессами самоорганизации.

При разработке Регламента роли саморегулирования было уделено отдельное внимание. В пояснительной записке приводятся четыре варианта, рассмотренные в процессе законотворчества: 1) введение добровольной системы маркировки; 2) секторальный *ad hoc* подход; 3) риск-ориентированный подход; 3+) риск-ориентированный подход, дополненный «кодексами поведения» для систем ИИ, не представляющих высокий риск; 4) обязательные требования для всех систем ИИ вне зависимости от представляемого ими риска. Особое внимание уделено кодексам поведения, т.е. создаваемым на уровне отдельных корпораций этическим нормам. Отметим большую гибкость такого подхода по сравнению, например, с четвертым вариантом, предполагающим введение обязательных требований для всех систем ИИ, а также большую строгость и комплексность по сравнению с первой опцией – введением процедуры добровольной маркировки продукции, и вторым вариантом – секторальным подходом.

5) *Динамичность и неравновесность, способность переходить от длительного состояния равновесия к резким изменениям.*

Средне- и долгосрочные эффекты от масштабного и интенсивного внедрения ИИ-технологий во все сферы жизни общества до сих пор подробно не изучены. Законодатели ЕС отдельно отметили в пояснительной записке, что при оценке различных вариантов регулирования систем ИИ учитывались возможные социально-экономические последствия и контексты каждого из них. ЕС также уделяет внимание научным исследованиям в области ИИ, в том числе для того, чтобы составить обоснованные прогнозы возможных изменений в смежных сферах.

6) *Междисциплинарность* объяснительных подходов теории сложных систем. Ее часто привлекают для объяснения социальных, политических и иных процессов, в исследованиях международных отношений. Однако П. Кейрни [Cairney, 2012] отмечает, что даже самые сложные исследования, использующие математические

методы моделирования, часто получают от практиков упреки за абстрактность и некоторую очевидность рекомендаций: «сделать модель управления более гибкой и адаптивной», «разработать новый инструментарий», «более комплексно рассматривать контексты» и т.д. В связи с этим необходимо дополнить анализ представленной в Регламенте модели, основываясь на теории многоуровневого управления.

### ***Взаимодействие акторов ИИ: попытка выстраивания многоуровневой системы***

Оценить предлагаемую в Регламенте модель можно с использованием теории, применяемой для рассмотрения взаимодействия институтов ЕС, а в последние годы и для формирования его цифровой политики. Это *теория многоуровневого управления*. Она позволяет перевести разговор об адаптивности системы в более практическую плоскость, не отрицая динамического характера взаимодействия системы и ее элементов как между собой, так и с внешней средой, и встроить разговор о регулировании ИИ в более широкий контекст политики ЕС в отношении цифровых технологий.

Цифровая политика ЕС через призму теории многоуровневого и экспериментального управления является объектом исследования М. Маркут. Экспериментальное управление возникает в условиях неясности целей и предполагает, что эти цели вырабатываются совместно в процессе переговоров между институтами или несколькими акторами, которые выступают как несколько отдельных центров принятия решений. Поскольку акторы не состоят в иерархической связи друг с другом, для разрешения противоречий задействуются неформальные механизмы. Такая система функционирует только при надлежащем взаимодействии наднационального и национального уровней [Marcut, 2020]. В этом смысле показательна проанализированная выше многоступенчатая разработка этических норм для ИИ.

Теория многоуровневого управления, в свою очередь, предполагает привлечение большего количества акторов к процессу управления для достижения лучших результатов. В ней большее внимание уделяется наднациональным институтам, которые не агрегируют предпочтения государств, а выступают независимой структурой, вырабатывающей и направляющей политику. М. Маркут отмечает, что в определении цифровой политики руководящая и координирующая роль все еще принадлежит Еврокомиссии, хотя в последнее время просматривается тенденция к большей демократизации. В частности, в вопросе регулирования ИИ учитывалось мнение Европарламента.

Обратимся к предлагаемой модели взаимодействия основных акторов ИИ. Ряд полномочий в текущей редакции документа делегируется на национальный уровень, когда государства смогут применять нормы с учетом своего законодательства, а контроль за соблюдением Регламента будут осуществлять национальные компетентные органы.

Предполагается, что координирующим органом станет новая структура – Европейский совет по ИИ, основные функции которого – обеспечение эффективного сотрудничества национальных регуляторов и Еврокомиссии, а также помощь и ре-

комендации для наднациональных органов. Совет сможет приглашать на заседания внешних экспертов и наблюдателей, взаимодействовать с заинтересованными игроками и в должной мере сообщать о своих действиях. При этом ЕК сохранит влияние на Совет, председательствуя в нем.

Отдельное место в системе взаимодействия отводится механизмам саморегулирования: согласно Регламенту, страны ЕС и Еврокомиссия будут поощрять разработку и применение кодексов поведения, авторами которых смогут выступить индивидуальные поставщики услуг и организации, выражающие их интересы. Таким образом, в Регламенте выстроена многоуровневая система взаимодействия, где направляющей и руководящей силой остается ЕК.

В регулировании ИИ в Евросоюзе просматриваются две тенденции: во-первых, тяготение к научному обеспечению разработки программных документов, политики и норм регулирования в области ИИ; во-вторых – постепенный переход от мягких форм регулирования к более обязывающим.

### **Выводы**

ЕС стоит перед серьезным вызовом регулирования быстроразвивающейся технологии ИИ, намереваясь применять ее в самых широких сферах. ЕС приходится сочетать выполнение двух задач: регулирование ИИ и стимулирование развития самой технологии. Сейчас можно выделить два основных вектора деятельности ЕС в сфере регулирования ИИ: этический («мягкий») и нормативно-правовой («жесткий»).

Подход ЕС к технологиям ИИ можно охарактеризовать как комплексный. В документах обоснована необходимость широкомасштабного внедрения ИИ, подчеркнуты его мощный инновационный импульс и экономический потенциал. Одновременно обозначены глобальные вызовы и риски, связанные с качественными изменениями в экономической и в социо-гуманитарной сфере в результате интенсивного внедрения ИИ, а также необходимость минимизации возможных пагубных воздействий ИИ в средне- и долгосрочной перспективе и сохранения траектории его развития.

На фоне очевидного отставания нормативно-правовых норм от динамично изменяющихся реалий, с одной стороны, и опасений зарегулировать и затормозить развитие столь инновационной сферы как ИИ, с другой, в европейских странах, начиная с 2017 г., проявляется тенденция к самоорганизации.

ЕС делает ставку на развитие концепции «доверенного ИИ», которая предполагает соответствие систем ИИ семи этическим требованиям: сохранение акторности и контроля человека; техническая надежность и безопасность; защита частной жизни и управление данными; прозрачность; разнообразие, недискриминация и справедливость; благосостояние окружающей среды и общества и подотчетность. Эти требования были зафиксированы в документе мягкого права – Руководящих принципах по этике для «доверенного» ИИ (2019).

В 2021 г. ЕК представила первый проект всеобъемлющего законодательного документа в сфере ИИ – Регламента о гармонизации правил, применяемых к систе-

мам ИИ. Документ выстраивает многоуровневую систему взаимодействий акторов ИИ при сохранении направляющей и руководящей роли Еврокомиссии. Реальный эффект документа в случае его принятия проявится не сразу, и многое будет зависеть от воплощения его норм на практике. Однако стоит отметить, что более простые решения, которые были удобнее с точки зрения контроля (введение обязательных требований ко всем системам ИИ), или же, напротив, перекладывали больше обязанностей на участников рынка (введение добровольной сертификации), были отвергнуты в процессе работы над законопроектом и общественных консультаций. Это – наряду с тенденцией к постоянному вовлечению экспертов в разработку политических рекомендаций, инструментов, методик оценки систем ИИ и дифференциации требований и подходов к различным группам акторов и систем ИИ – говорит о том, что законодатели в ЕС пришли к пониманию технологической политики в отношении ИИ как сложной системы.

Развитие компетенций ЕС и потенциальное лидерство в этой сфере, в том числе в регулировании ИИ, укрепляет глобальную роль ЕС как актора. В этой конфигурации ЕС стремится занять позицию разработчика правил (*rule-maker*), вырабатывая нормы поведения.

Право имеет четко ограниченную сферу применения, в то время как мораль подобных ограничений, национальных границ применения, не имеет. Приверженность этической парадигме регулирования повышает вес ЕС на международной арене, делая этику регулирования ИИ одним из компонентов нормативной силы.

### СПИСОК ЛИТЕРАТУРЫ

Базаркина Д.Ю. (2021) Эволюция подходов к противодействию гибридным угрозам в стратегическом планировании Европейского союза. *Современная Европа*, № 6. С. 133–143. DOI: <http://dx.doi.org/10.15211/soveurope62021133143>

Доверенный искусственный интеллект: концепция и документы. *Этика и «цифра»: от проблем к решениям* (2021) Под ред. Е.Г. Потаповой, М.С. Шклярук. РАНХиГС, Москва. С. 90–122.

Истомин И.А. (2021) *Логика поведения государств в международной политике*. Аспект Пресс, Москва. 88 с.

Истомин И.А., Дегтярев Д.А. (2015) Системное моделирование международных отношений *Мировая экономика и международные отношения*, № 11. С. 17–30.

Незнамов А.В. (2019) Правовые аспекты реализации Национальной стратегии развития искусственного интеллекта до 2030 года. *Вестник университета имени О.Е. Кутафина (МГЮА)*, № 12. С. 82–88.

Потёмкина О.Ю. (2019) Лучше, чем люди? Политика ЕС в области искусственного интеллекта. *Научно-аналитический вестник ИЕ РАН*, № 5. С. 16–21. DOI: <http://dx.doi.org/10.15211/vestnikieran520191621>

Федоров М.В., Цветков Ю.Д. (2020) Этика искусственного интеллекта в деятельности ЮНЕСКО. *РСМД*. URL: <https://russiancouncil.ru/analytics-and-comments/analytics/etika-iskusstvennogo-intellekta-v-deyatelnosti-yunesko-voprosy-politiki-prava-i-perspektivy-ravnopra/> (дата обращения: 17.05.2022)

Цыганков П.А. (2003) *Теория международных отношений*. Гардарики, Москва. 590 с.

Bazarkina D, Pashentsev E. (2019) Artificial Intelligence and New Threats to International Psychological Security. *Russia in Global Affairs*, 1.

Burns S., Bradley J., Powell M. (2021) *A robust approach? EU proposes framework for mandatory regulation of AI*. Lexology.com. URL: [https://www.lexology.com/library/detail.aspx?g=d85fe0cd-6f91-44a4-82d5-a5d5217a6668#:~:text=Questions%3F-](https://www.lexology.com/library/detail.aspx?g=d85fe0cd-6f91-44a4-82d5-a5d5217a6668#:~:text=Questions%3F-,A%20robust%20approach%3F,for%20mandatory%20regulation%20of%20AI&text=On%2021%20April%202021%2C%20the,legal%20framework%20for%20AI%20globally)

,A%20robust%20approach%3F,for%20mandatory%20regulation%20of%20AI&text=On%2021%20April%202021%2C%20the,legal%20framework%20for%20AI%20globally (дата обращения: 17.05.2022)

Cairney P. (2012) Complexity Theory in Political Science and Public Policy. *Political Studies Review*. Vol. 10, pp. 346–358.

Ethics Guidelines for Trustworthy AI (2019). *European Commission*. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (дата обращения: 17.05.2022)

Gutierrez C.I. and Marchant G.E. (2021) A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence. *SSRN*. URL: <http://dx.doi.org/10.2139/ssrn.3855171>.

Hoffmann S. (1981) *Duties Beyond Borders: On the Limits and Possibilities of Ethical International Politics*. N.Y. 46 p.

Kuleshov A., Ignatiev A., Abramova A., Marshalko G. (2020) Addressing AI Ethics Through Codification. *Proceedings of the 2020 International Conference Engineering Technologies and Computer Science (EnT)*. 24-26 June 2020. URL: <https://ieeexplore.ieee.org/document/9140627> (дата обращения: 17.05.2022)

Marcut M. (2020) *The Governance of Digital Policies. Towards a New Competence in the European Union*. Palgrave Macmillan. 136 p.

Mueller B. (2021) How Much Will the Artificial Intelligence Act Cost Europe? *Center for Data Innovation*, 16 p. URL: <https://www2.datainnovation.org/2021-aia-costs.pdf> (дата обращения: 17.05.2022)

Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (2021). *European Commission*. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206> (дата обращения: 17.05.2022)

White Paper on Artificial Intelligence. A European Approach to Excellence and Trust (2020). *European Commission*. URL: [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf) (дата обращения: 17.05.2022)

## **The Dialectics of Ethics and Law in the Regulation of Artificial Intelligence: Case of the EU**

**E.A. Maslova\***

*Candidate of Sciences (Politics)*

*Associate Professor and Senior Research Fellow at MGIMO University*

*76, Vernadsky prospect, Moscow, Russia, 119454*

*Senior Research Fellow at the Institute of Europe of the Russian Academy of Sciences. 11-3*

*Mokhovaya, Moscow, Russia, 125009*

**\*E-mail:** [e.maslova@inno.mgimo.ru](mailto:e.maslova@inno.mgimo.ru)

**E.D. Sorokova\*\***

*PhD Student, MGIMO University*

*76, Vernadsky prospect, Moscow, Russia, 119454*

**\*\*E-mail:** [sorokova.e.d@my.mgimo.ru](mailto:sorokova.e.d@my.mgimo.ru)

**Abstract.** The European Union is determined to achieve leading positions in the development of artificial intelligence (AI) systems, which constitutes a whole direction of its activities in this sphere. At the same time, the introduction of AI technologies has an outstanding transformative potential, which poses a challenge not only for global development, but also for the European bureaucracy that has faced the dilemma of AI regulation. Its essence is to build a management model that would minimise uncertainty and risks of catastrophic consequences for the state, society and the individual without holding back technological progress. The article explores the European approach to AI regulation, both in the ethical and legal fields. The authors consider the main ethical principles that EU has set in the field of AI, as well as their evolution, and analyse the emergent approach to the legal regulation of AI using the theories of complex systems and multilevel governance. It is concluded that the EU prioritises the development of "trustworthy AI", wherein a gradual transition from 'soft law' to more binding forms of regulation can be traced. Another distinctive feature of AI regulation in the EU is the consecutive scientific support in the development of policy papers, reports and regulations.

**Key words:** ethics of international relations, AI ethics, artificial intelligence, AI, European Union, complex systems.

**DOI:** 10.31857/S0201708322050023

**EDN:** glofsa

## REFERENCES

Bazarkina D.Yu. (2021) Evolyuciya podhodov k protivodejstviyu gibrnidnym ugrozam v strategicheskom planirovanii Evropejskogo soyuza [Evolving Approaches to Countering Hybrid Threats in European Union Strategic Planning]. *Sovremennaya Evropa*, No 6, pp. 133–143. DOI: <http://dx.doi.org/10.15211/soveurope62021133143> (In Russian).

Bazarkina D, Pashentsev E. (2019) Artificial Intelligence and New Threats to International Psychological Security. *Russia in Global Affairs*, 1.

Burns S., Bradley J., Powell M. (2021) *A robust approach? EU proposes framework for mandatory regulation of AI*. Lexology.com. URL: <https://www.lexology.com/library/detail.aspx?g=d85fe0cd-6f91-44a4-82d5-a5d5217a6668#:~:text=Questions%3F-,A%20robust%20approach%3F,for%20mandatory%20regulation%20of%20AI&text=On%2021%20April%202021%2C%20the,legal%20framework%20for%20AI%20globally> (accesses: 17.05.2022)

Cairney P. (2012) Complexity Theory in Political Science and Public Policy. *Political Studies Review*. Vol. 10, pp. 346–358.

Cygankov P.A. (2003) *Teoriya mezhdunarodnyx otnoshenij* [Theory of International relations]. Gardariki, Moscow. 590 p. (In Russian).

*Doverennyj iskusstvennyj intellekt: koncepciya i dokumenty. Etika i «cifra»: ot problem k resheniyam* [Trusted Artificial Intelligence: concept and documents. Ethics and "digital": from problems to solutions] (2021). Pod red. E.G. Potapovoj, M.S. Shklyaruk. RANHiGS, Moscow. 184 p. (In Russian).

- Ethics Guidelines for Trustworthy AI (2019). *European Commission*. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (accesses: 17.05.2022)
- Gutierrez C.I. and Marchant G.E. (2021) A Global Perspective of Soft Law Programs for the Governance of Artificial Intelligence. *SSRN*. URL: <http://dx.doi.org/10.2139/ssrn.3855171>.
- Hoffmann S. (1981) *Duties Beyond Borders: On the Limits and Possibilities of Ethical International Politics*. N.Y. 46 p.
- Istomin I.A. (2021) *Logika povedeniya gosudarstv v mezhdunarodnoj politike* [The Logic of States' Actions in World Politics]. Aspekt Press, Moscow. 88 p. (In Russian).
- Istomin I.A., Degrtarev D.A. (2015) Sistemnoe modelirovanie mezhdunarodnyh otnoshenij [Systemic Modelling of International Relations]. *World Economics and International Relations*. No 11, pp. 17–30. (In Russian).
- Kuleshov A., Ignatiev A., Abramova A., Marshalko G. (2020) Addressing AI Ethics Through Codification. *Proceedings of the 2020 International Conference Engineering Technologies and Computer Science (EnT)*. 24–26 June 2020. URL: <https://ieeexplore.ieee.org/document/9140627> (accesses: 17.05.2022)
- Marcut M. (2020) *The Governance of Digital Policies. Towards a New Competence in the European Union*. Palgrave Macmillan. 136 p.
- Mueller B. (2021) How Much Will the Artificial Intelligence Act Cost Europe? *Center for Data Innovation*, 16 p. URL: <https://www2.datainnovation.org/2021-aia-costs.pdf> (accesses: 17.05.2022)
- Neznamov A.V. (2019) Pravovye aspekty realizacii Nacional'noj strategii razvitiya iskusstvennogo intellekta do 2030 goda [Legal aspects of the implementation of the National Strategy for the Development of Artificial Intelligence until 2030]. *Vestnik universiteta imeni O.E. Kutafina*, No 12, pp. 82–88. (In Russian).
- Potemkina O.Yu. (2019) Luchshe, chem lyudi? Politika ES v oblasti iskusstvennogo intellekta [Better than humans? EU Artificial Intelligence Policy]. *Nauchno-analiticheskij vestnik IE RAN*. No 5, pp. 16–21. DOI: <http://dx.doi.org/10.15211/vestnikieran520191621> (In Russian).
- Fedorov M.V., Tsvetkov Yu.D. (2020) Etika iskusstvennogo intellekta v deyatelnosti YUNESKO [The Ethics of Artificial Intelligence in the Work of UNESCO]. *RIAC*. URL: <https://russiancouncil.ru/analytics-and-comments/analytics/etika-iskusstvennogo-intellekta-v-deyatelnosti-yunesko-voprosy-politiki-prava-i-perspektivy-ravnopra/> (accesses: 17.05.2022) (In Russian).
- Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts (2021). *European Commission*. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206> (accesses: 17.05.2022)
- White Paper on Artificial Intelligence. A European Approach to Excellence and Trust (2020). *European Commission*. URL: [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf) (accesses: 17.05.2022)