

Секция «Международная безопасность: новые и традиционные вызовы и угрозы»

Геополитические вызовы и перспективы международного сотрудничества в условиях разработки этического регулирования ИИ.

Научный руководитель – Рамич Мирзет Сафетович

Потемкина Виктория Олеговна

Студент (бакалавр)

Российский университет дружбы народов, Факультет гуманитарных и социальных наук,
Москва, Россия

E-mail: vika.pote465@mail.ru

В то время, как проблема предотвращения угроз политической нестабильности из-за неконтролируемого влияния ИИ остается общепризнанной, до сих пор существуют препятствия для достижения глобального единства по выработке стандартов и законов. Несмотря на попытки координации усилий, что подтверждается участием 28 стран на Международном саммите по безопасности ИИ в 2023 г., конечным решением является подписание «Декларации Блетчли» [14], которая не носит обязательный характер и не служит международным правовым договором. [3]

Подход ЕС к этике ИИ заключается в том, что в основу заложена нормативно-правовая система, гарантирующая соблюдение прав человека, прозрачность и подотчетность систем ИИ: ещё в 2019 г. ЕС ввел понятие «ответственного ИИ» в законодательство [8], а в 2024 г. опубликовал первый документ, который имеет юридическую силу - «Регламент Европейского союза об ИИ» [12].

Говоря о сотрудничестве западных стран в сфере регулирования ИИ, то следует выделить, во-первых, Глобальное партнерство по ИИ (GPAI), с помощью которого продвигаются общие демократические ценности, а во-вторых, Совет по торговле и технологиям США-ЕС, который дает пространство для согласования практик применения ИИ. [1]

В силу того, что стратегические цели Соединенных Штатов в первую очередь направлены на достижение мирового лидерства в ИИ, то в вопросе регулирования технологий акцент смещен на минимизацию жестких ограничений, для чего они полагаются на саморегулирование частного сектора и корпоративную ответственность. Такой гибкий подход обеспечивает отсутствие сдерживания инноваций для создания технологического разрыва с КНР. [2]

Поскольку действие опубликованного в 2022 г. «Билля о правах в области ИИ» [5] ограничено только внутренними правовыми рамками США, то уже в 2023 г. был издан «Исполнительный указ о безопасной, надёжной и заслуживающей доверия разработке и использовании ИИ» [13], который помимо продвижения американских этических стандартов в ИИ также призывает к внедрению согласованных стандартов, сотрудничеству и обмену информацией в глобальном масштабе. Значимость активного участия в принятии международных стандартов, которые могли бы обеспечивать их соответствие американским интересам, в том числе выражается в обнародовании Национальным институтом стандартов и технологий в 2024 г. «Плана глобального взаимодействия по стандартам ИИ» [10].

В противовес западным подходам, КНР выступает с готовностью разрабатывать этические принципы ИИ, исходя из собственных политических приоритетов. Подобная система характеризуется централизованным регулированием, где за государством остается главная функция по осуществлению полного контроля над разработками и применением ИИ с целью обеспечения национальной безопасности и соблюдения социалистических ценностей. [4]

Разрабатывая собственные механизмы регулирования ИИ, Китай в качестве мировой альтернативы преследует цель экспорта своей модели. Так, в 2021 г. были выпущены важные документы, которые сфокусированы на обеспечении государственной безопасности, цифрового суверенитета и управлением информацией, что страна стремится продвигать приоритетно в развивающихся странах: «Китайские стандарты 2035» [6], «Закон Китайской Народной Республики о защите персональной информации» [9], «Регламенты по управлению алгоритмами рекомендаций для интернет-информационных сервисов» [11], «Этические нормы для нового поколения ИИ» [7].

Подводя итог, в условиях складывающейся разобщенности в разработке этики регулирования ИИ, странам нужно объединить усилия для предотвращения угрозы всей человеческой цивилизации в случае потери контроля над технологиями.

Источники и литература

- 1) Hine E. Governing Silicon Valley and Shenzhen: Assessing a New Era of Artificial Intelligence Governance in the United States and China. DISO, 2024. Vol. 3, No. 50. DOI: 10.1007/s44206-024-00138-7.
- 2) Hine E., Floridi L. Artificial Intelligence with American Values and Chinese Characteristics: A Comparative Analysis of American and Chinese Governmental AI Policies // AI & Society. 2024. Vol. 39. P. 257–278. DOI: 10.1007/s00146-022-01499-8.
- 3) [U+6C88] [U+9038], [U+9AD8] [U+745C]. [U+5927] [U+56FD] [U+7ADE] [U+4E89] [U+80CC] [U+6650+154-155. DOI: 10.13851/j.cnki.gjzw.202403003.
- 4) [U+4E01] [U+8FEA]. [U+8D1F] [U+8D23] [U+4EFB] [U+4EBA] [U+5DE5] [U+667A] [U+80FD] [U+FF1119.
- 5) Blueprint for an AI Bill of Rights // The White House Archive. URL: <https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/> (accessed: 03.02.2025).
- 6) China Standards 2035 // Government of the People's Republic of China. URL: <http://www.sac.gov.cn/> (accessed: 03.02.2025).
- 7) Ethical Norms for New Generation Artificial Intelligence // Government of the People's Republic of China. URL: <http://www.most.gov.cn/> (accessed: 03.02.2025).
- 8) Ethics Guidelines for Trustworthy AI // European Commission Digital Strategy. URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (accessed: 03.02.2025).
- 9) Personal Information Protection Law (PIPL) // Government of the People's Republic of China. URL: <https://www.npc.gov.cn/> (accessed: 03.02.2025).
- 10) Plan for Global Engagement on AI Standards // National Institute of Standards and Technology. URL: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-5.pdf> (accessed: 03.02.2025).
- 11) Provisions on the Management of Algorithmic Recommendations in Internet Information Services // Government of the People's Republic of China. URL: <https://www.miit.gov.cn/> (accessed: 03.02.2025).
- 12) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts // EUR-Lex. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689> (accessed: 03.02.2025).

- 13) Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence // Federal Register of the United States Government. URL: <https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence> (accessed: 03.02.2025).
- 14) The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023 // UK Government Official Publications. URL: <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023> (accessed: 03.02.2025).