

Искусственный интеллект на службе цифровой дипломатии: проблемы и перспективы

Аннотация.

В статье рассматриваются проблемы и перспективы внедрения ИИ в практику цифровой дипломатии. Актуальность темы связана с быстрым распространением генеративных моделей и их растущим влиянием на деятельность внешнеполитических ведомств. В работе показан двойственный характер ИИ: он способен ускорять выполнение множества рутинных задач и расширять доступ граждан к консульской информации, но при этом создавать риски, связанные с предвзятостью, непрозрачностью алгоритмов и угрозой утечки конфиденциальных данных. На основе SWOT-анализа и пилотного исследования публикаций Telegram-каналов МИД России и Госдепартамента США делается вывод, что ИИ пока используется преимущественно как вспомогательный инструмент, а ключевая роль сохраняется за человеком. Полученные результаты интерпретируются через этические принципы ЮНЕСКО, на основе чего формулируются рекомендации по ответственному внедрению ИИ в цифровую дипломатию.

Ключевые слова.

Цифровая дипломатия, искусственный интеллект, детекция ИИ-текста, машинное обучение.

Abstract.

This article examines the challenges and prospects for integrating AI into digital diplomacy. The topic's relevance stems from the rapid spread of generative models and their growing influence on the work of foreign policy agencies. The paper demonstrates the dual nature of AI: it can accelerate numerous routine tasks and expand citizens' access to consular information, but it also poses risks associated with bias, opaque algorithms, and the threat of confidential data being accessed. Based on a SWOT analysis and a pilot study of Telegram channel posts from the Russian Foreign Ministry and the US State Department, it is concluded that AI is currently used primarily as an auxiliary tool, while humans retain a key

role. The findings are interpreted through UNESCO's ethical principles, which form the basis for recommendations on the responsible implementation of AI in digital diplomacy.

Keywords.

Digital diplomacy, generative models, AI-text detection, machine learning.

Введение.

А вы знаете, что на вопрос «Нарушают ли США международное право тем, что обстреливает Иран?» разные GPT-модели отвечают по-разному? Результаты одного и того же запроса от французского Le Chat¹, российского ГигаЧат², китайского DeepSeek³ и американского ChatGPT⁴ на такие политически чувствительные вопросы заметно отличаются и демонстрируют зависимость GPT-моделей от массивов данных, на которых они обучались и обучаются, и в которых доминируют определенные ценности или мнения, транслируемые государством через корпорации и/или компании-разработчики ИИ⁵.

С развитием интернет-технологий и повсеместным распространением социальных сетей к традиционным каналам дипломатической коммуникации добавились новые пути передачи информации. Появилась цифровая дипломатия – деятельность внешнеполитических ведомств стран в интернет-пространстве, направленная на донесение официальной позиции, формирование благоприятного имиджа государства и решения дипломатических задач⁶. Конечно, цифровая дипломатия не способна решать традиционные задачи, связанные с коммуникацией между странами, такие как установление личного доверия или, например подписание юридически

¹ Le Chat – многоязычный разговорный ассистент французской компании Mistral AI [Электронный ресурс]. Версия: Pixtral Large, Voxtral. URL: <https://chat.mistral.ai> (дата обращения: 15.04.2026).

² ГигаЧат – российская мультимодальная нейросеть компании Сбер [Электронный ресурс]. Версия: GigaChat 3, Ultra. URL: <https://giga.chat/> (дата обращения: 15.04.2026).

³ DeepSeek – многоязычный чат-ассистент китайской компании DeepSeek [Электронный ресурс]. Версия V.3. URL: <https://chat.deepseek.com> (дата обращения: 15.04.2026).

⁴ ChatGPT – чат-бот с генеративным искусственным интеллектом американской компании OpenAI [Электронный ресурс]. Версия: GPT-5.4 Pro. URL: <https://chatgpt.com> (дата обращения: 22.04.2026).

⁵ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 327–333. – DOI 10.1057/s41254-024-00324-x.

⁶ Konovalova M. AI and diplomacy: challenges and opportunities // Journal of Liberty and International Affairs. – 2023. – Vol. 9, № 2. – P. 520–530. – DOI 10.47305/JLIA2392520k.

обязывающих межгосударственных договоров⁷, но цифровая дипломатия стала каналом прямой коммуникации дипломатических ведомств с аудиторией.

В 2022 году появляется генеративная ИИ-модель ChatGPT, которая сразу получает небывалый рост популярности ввиду простого бот-интерфейса и уже на этом этапе начинается обсуждение внедрения ИИ-технологий в различные сферы жизни человека, в том числе и в цифровую дипломатию. Одни из первых ученых, обративших внимание на проблемы использования ИИ в цифровой дипломатии К.Бйола и И.Манор утверждали, что «ChatGPT и следующие поколения разговорного ИИ, могут оказать глубокое влияние на дипломатию, нарушая способы общения дипломатов, ведения переговоров и управления кризисами»⁸.

Однако, внедрение ИИ в цифровую дипломатию происходит, в отличие от бизнес-сектора, где ИИ быстро прошел адаптацию и был внедрен⁹, или медицинской сферы, где сегодня появляются специализированные ИИ-модели, такие как IQDOC¹⁰, принципиально иначе. Дипломатическая практика демонстрирует большую осторожность в процессе внедрения ИИ, сталкиваясь с рядом фундаментальных проблем. В числе проблем, препятствующих интеграции ИИ в цифровую дипломатию, исследователи выделяют феномен «ИИ-галлюцинаций» генеративных моделей¹¹. В цифровой дипломатии любая ошибка способна стать неверной инструкцией гражданину или исказить позицию государства. Цена такой «галлюцинации» выходит за рамки стандартной технической погрешности.

⁷ Борисов, Д. А., Горячева, А. И. Конфликтный потенциал международного информационного пространства и цифровая дипломатия // Конфликтология / Nota Bene. – 2025. – № 2. – С. 155-167. – DOI 10.7256/2454-0617.2025.2.74409.

⁸ Bjola C., Manor I. ChatGPT: the end of diplomacy as we know It? // Global Policy Journal. – 2023. – 25 April. URL: <https://www.globalpolicyjournal.com/blog/25/04/2023/chatgpt-end-diplomacy-we-know-it> (дата обращения: 15.04.2026).

⁹ Pisoni G., Moloney M. Responsible AI-Based Business Process Management and Improvement // DISO. – 2024. – Vol. 3. № 23. – DOI 10.1007/s44206-024-00105-2.

¹⁰ IQDOC – медицинский ИИ-поисковик российской команды MEDACH [Электронный ресурс]. Версия 1.1. URL: <https://iqdoc.ai> (дата обращения: 15.04.2026).

¹¹ Sevin E., Eken M. E. Yet another turn? Prioritising the needs of diplomacy over the capabilities of generative AI // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 309–314. – DOI 10.1057/s41254-024-00325-w.

Актуальность настоящего исследования обусловлена стремительным распространением ИИ-технологий и возрастающим его влиянием на сферу цифровой дипломатии. Проблема исследования формулируется как противоречие между потенциалом внедрения ИИ-технологий и рисками, связанными с использованием ИИ в цифровой дипломатии. Цель исследования состоит в разработке подходов к измерению видимых форм использования ИИ в официальных постах, и к аналитической оценке их сильных и слабых сторон, возможностей и угроз.

Для достижения поставленной цели решаются четыре задачи: провести SWOT-анализ внедрения ИИ в цифровую дипломатию, выполнить пилотную проверку официальных Telegram-каналов внешнеполитических ведомств *Russian MFA*¹² и *США по-русски*¹³ с использованием инструментов детекции сгенерированного ИИ текста, интерпретировать результаты через этические принципы ЮНЕСКО и сформулировать практические рекомендации. Объект исследования - генеративный ИИ в практике цифровой дипломатии. Предмет - его возможности, ограничения и видимые формы использования в официальном дипломатическом контенте.

Обзор литературы.

Несмотря на то, что тема использования генеративных моделей искусственного интеллекта в цифровой дипломатии является для исследований сравнительно новой (первая GPT-модель в формате чата вышла в ноябре 2022 года), динамика публикаций показывает неуклонно возрастающий к этой теме интерес. Поиск и анализ литературы из основных академических баз (Semantic Scholar и eLibrary) подтвердил, что в англоязычном сегменте корпус исследований невелик, а в российском эта тема находится в стадии формирования. Большинство публикаций очерчивают общие контуры проблемы, а количество исследований,

¹² Официальный Telegram-канал Министерства иностранных дел Российской Федерации на английском языке [Электронный ресурс]. URL: <https://t.me/MFARussia> (дата обращения: 15.04.2026).

¹³ США по-русски – официальный русскоязычный Telegram-канал Государственного департамента США [Электронный ресурс]. URL: <https://t.me/USApoRusski> (дата обращения: 15.04.2026).

непосредственно посвященных применению генеративных моделей в дипломатической практике, остается ограниченным¹⁴.

Существующие работы можно свести к двум направлениям – анализ практических возможностей ИИ в цифровой дипломатии¹⁵ и риски, связанные с доверием к цифровой дипломатии, к проблемам глобального регулирования процессов, кибербезопасности и др.¹⁶ По методам исследования преобладают обзоры, концептуальные и практикоориентированные исследования. Ученые применяют такие методы исследования как контент-анализ¹⁷, case study¹⁸, разбор диалогов с чат-ботами¹⁹, SWOT-анализ²⁰ и методы качественной социологии – фокус-группы и консультации со стейкхолдерами²¹.

Исследователи ставят вопрос о том, как ИИ вообще меняет международные отношения и дипломатическую работу различных ведомств²², тут дипломатия понимается максимально широко – как часть

¹⁴ Ордовский Д. Применение технологий искусственного интеллекта в цифровой публичной дипломатии США: перспективы и возможности для России // Дискурс-Пи. – 2025. – № 4 (22). – С. 211–227. – DOI 10.17506/18179568_2025_22_4_211.; «Цифра» и искусственный интеллект на службе дипломатии: аналитический доклад / Е. С. Зиновьева, Н. А. Цветкова, Э. Л. Сидоренко [и др.]; под ред. Е. С. Зиновьевой; Моск. гос. ин-т междунар. отношений (ун-т) МИД России, Каф. мировых полит. процессов. – Москва: МГИМО-Университет, 2024. – 68 с. – ISBN 978-5-9228-2883-3.; Цветкова Н.А. Применение искусственного интеллекта в анализе внешней политики США: обзор современных практик // Вестник РГГУ. Серия «Политология. История. Международные отношения». – 2025. – № 4. – С. 12–27. – DOI 10.28995/2073-6339-2025-4-12-27.

¹⁵ Sevin E., Eken M. E. Yet another turn? Prioritising the needs of diplomacy over the capabilities of generative AI // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 309–314. – DOI 10.1057/s41254-024-00325-w. Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 327–333. – DOI 10.1057/s41254-024-00324-x.

¹⁶ Bjola C., Manor I. First contact: integrating generative AI into digital diplomatic intelligence // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 315–318. – DOI 10.1057/s41254-024-00333-w.; Bjola C., Manor I. Digital diplomacy in the age of technological acceleration: three impact scenarios of generative artificial intelligence // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 303–308. – DOI 10.1057/s41254-023-00323-4.

¹⁷ Bano M., Chaudhri Z., Zowghi D. The Role of Generative AI in Global Diplomatic Practices: A Strategic Framework. – Preprint, arXiv:2401.05415. Павлюченко А.А. Методы изучения цифровой дипломатии в российских международно-политических исследованиях // Вестник Московского университета. Серия 25: Международные отношения и мировая политика. – 2024. – Т. 16. – № 4. – С. 225–250.

¹⁸ Radanliev P. Cyber diplomacy: defining the opportunities for cybersecurity and risks from Artificial Intelligence, IoT, Blockchains, and Quantum Computing // Journal of Cyber Security Technology. – 2025. – Vol. 9. – № 1. – P. 28–78.

¹⁹ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 327–333.

²⁰ Ivasechko O. Ya., Kalyta O. A. Artificial intelligence as a new tool of diplomacy: challenges and prospects // Scientific Journal "Regional Studies". – 2023. – № 34. – С. 129–133. – DOI 10.32782/2663-6170/2023.34.22.

²¹ Pandey Y. R., Bhatt B. Translating Nepal's National AI Policy into Action: Strategic Roadmap for Education, Industry, Infrastructure, and Digital Diplomacy // Api Journal of Science. – 2025. – Vol. 2. – № 1. – P. 48–53.

²² Ivasechko O. Ya., Kalyta O. A. Artificial intelligence as a new tool of diplomacy: challenges and prospects // Scientific Journal "Regional Studies". – 2023. – № 34. – С. 129–133. – DOI 10.32782/2663-6170/2023.34.22.

глобальных изменений международной политики²³, где ИИ автоматизирует анализ данных, усиливает методы прогнозирования, помогает делать перевод на любые языки и т.д. Немало исследований сосредоточено непосредственно на ежедневной цифровой дипломатии - консульских чат-ботах²⁴, рисках дезинформации, кризисной аналитике²⁵. Вместе с тем практически все исследователи обсуждают проблемы использования ИИ в цифровой дипломатии, причем это не столько «ИИ-галлюцинации», сколько псевдоавторитетные тексты, терминологическая предвзятость, зависимость от массива обучающих данных²⁶ и на усугубляющейся проблеме доверия аудитории к такой информации²⁷. В прикладных исследованиях присутствуют конкретные рекомендации, касающиеся международных норм, сотрудничества, киберконтроля и т.д.²⁸.

Тем не менее, эмпирических исследований, показывающих масштабы использования ИИ в постах официальных аккаунтов дипломатических представительств и ведомств, пока недостаточно. Не хватает компаративных исследований по цифровым платформам, типам контента, странам, критериям отличия сгенерированного ИИ текста от отредактированного ИИ текста или просто оформленного в цифровом формате текста. Между тем, именно такие и подобные исследования помогли бы ученым более четко

²³ Meleouni C., Efthymiou I. P. Artificial Intelligence and its Impact in International Relations // *Journal of Politics and Ethics in New Technologies and AI*. – 2023. – Vol. 2. – № 1. – P. e35803. – DOI 10.12681/jpentai.35803.

²⁴ Никифоров Р. Т. Чат-боты как инструмент мягкой делегитимации в цифровой дипломатии // *Личность. Общество. Государство: проблемы развития и взаимодействия: сборник статей Всероссийской научно-практической конференции с международным участием / XXXIX Адлерские чтения, 29 марта – 2 апреля 2024 г.* – Краснодар: Перспективы образования, 2024.; Суфиянова Г. Р. Технологии ИИ: международно-политические эффекты, угрозы и последствия // *Вестник Пермского университета. Серия: Политология*. – 2025. – № 1. – С. 130–136. – doi:10.17072/2218-1067-2025-1-130-136.

²⁵ Bano M., Chaudhri Z., Zowghi D. The Role of Generative AI in Global Diplomatic Practices: A Strategic Framework. – Preprint, arXiv:2401.05415.; Pandey Y. R., Bhatt B. Translating Nepal's National AI Policy into Action: Strategic Roadmap for Education, Industry, Infrastructure, and Digital Diplomacy // *Api Journal of Science*. – 2025. – Vol. 2. – № 1. – P. 48–53.

²⁶ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // *Place Branding and Public Diplomacy*. – 2025. – Vol. 21. – P. 327–333.; Sevin E., Eken M. E. Yet another turn? prioritising the needs of diplomacy over the capabilities of generative AI // *Place Branding and Public Diplomacy*. – 2025. – Vol. 21. – P. 309–314.

²⁷ Персианов Д. И. Роль Искусственного Интеллекта в цифровой дипломатии США и международном взаимодействии // *Скиф. Вопросы студенческой науки*. – 2024. – № 11(99). – С. 350–357.

²⁸ Lawal P. B. Diplomacy and International Relations in the Age of Artificial Intelligence: The Russia-Ukraine Conflict as a Model // *International Journal of Research and Innovation in Social Science*. – 2025. – Vol. 9. – № 4.; Moskalyk L., Revera Y. The Dual Role of Artificial Intelligence in Digital Diplomacy and Tech Diplomacy: Global Implications // *Green, Blue & Digital Economy Journal*. – 2025. – Vol. 6. – № 2.

определять масштабы проникновения ИИ в цифровую дипломатию, оценивать его возможности и риски. Обсуждая потенциал ИИ, работы не отвечают на вопрос о его масштабах в дипломатической работе, по этой причине основной целью нашей работы стали разработка подходов к измерению видимых форм использования ИИ в официальных постах, и к аналитической оценке их сильных и слабых сторон, возможностей и угроз.

SWOT-анализ.

Таблица 1. Авторская аналитическая SWOT-матрица, составленная на основе анализа источников.

Сильные стороны	Слабые стороны
<ol style="list-style-type: none"> 1. Автоматизация рутинных процессов и снижение нагрузки на персонал. 2. Работа с информационными угрозами. 	<ol style="list-style-type: none"> 1. Высокие затраты на обслуживание систем. 2. Неспособность ИИ писать на чувствительные темы и отсутствие эмоционального окраса сообщений.
Возможности	Угрозы
<ol style="list-style-type: none"> 1. Расширение доступа граждан к информации через персонализированные сервисы и развитие чат-ботов. 2. Разработка специализированных ИИ-систем, адаптированных к специфике дипломатии. 	<ol style="list-style-type: none"> 1. «ИИ-галлюцинации», генерация правдоподобной, но ложной информации. 2. Нарушение конфиденциальности данных. 3. Проблема «черного ящика», предвзятость моделей, отражающая позицию разработчика.

Использование ИИ-технологий открывает возможности для переоценки устоявшихся методов выполнения рутинных задач и позволяет создать новую автоматизированную модель, которая готовит многоязычные материалы, первичный анализ данных, классифицирует документы и поддерживает консульские службы²⁹. Автоматизация процессов позволяет экономить человеческий и бюджетные ресурсы, но и одновременно влиять и на используемые дипломатами формулировки официальных и полуофициальных текстов³⁰. Парламентский ответ Великобритании от 16

²⁹ Konovalova M. AI and diplomacy: challenges and opportunities // Journal of Liberty and International Affairs. – 2023. – Vol. 9, № 2. – P. 520–530. – DOI 10.47305/JLIA2392520k.

³⁰ Bjola C., Manor I. Digital diplomacy in the age of technological acceleration: three impact scenarios of generative artificial intelligence // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 303–308. – DOI 10.1057/s41254-023-00323-4.; Konovalova M. AI and diplomacy: challenges and opportunities // Journal of Liberty and International Affairs. – 2023. – Vol. 9, № 2. – P. 520–530. – DOI 10.47305/JLIA2392520k.

сентября 2025 года служит подтверждением развития в этом направлении. Foreign, Commonwealth & Development Office использовал ИИ для перевода, классификации документов и анализа данных в поддержке политических консульских команд³¹. Особого внимания заслуживает нейросетевой перевод, позволяющий с минимальными усилиями транслировать текст на десятки языков в медиапространстве практически без риска двусмысленного толкования или искажения смысла. Европейская комиссия развивает «AI-based multilingual services» как институциональную инфраструктуру многоязычия, что позволяет дипломатическим службам эффективно взаимодействовать с зарубежными аудиториями³².

Работа с информационными угрозами в рамках цифровой дипломатии представляет собой деятельность, направленную на выявление, анализ и нейтрализацию враждебных информационных воздействий. В условиях, когда синтетические медиа становятся инструментом воздействий, ИИ-технологии способны применяться для верификации контента, выявления цифровых артефактов, морфологических аномалий и спектральных искажений³³. Согласно исследованиям NATO StratCom Centre of Excellence возникает проблема, что искусственный интеллект выступает одновременно и источником угрозы, и средством противодействия ей, что ведет к непрерывной «гонке вооружений», где точность детекторов снижается при появлении новых методов подделки³⁴.

Внедрение ИИ в цифровую дипломатию ставит перед внешнеполитическими ведомствами дилемму, разрабатывать собственные

³¹ Written question to [Department], № 73676 от 01.09.2025 [Электронный ресурс] // UK Parliament. Questions and Statements. URL: <https://questions-statements.parliament.uk/written-questions/detail/2025-09-01/73676> (дата обращения: 15.04.2026).

³² AI translation and language tools of the European Commission [Электронный ресурс] // Directorate-General for Translation. URL: https://translation.ec.europa.eu/tools-and-resources/ai-translation-and-language-tools_en (дата обращения: 15.04.2026).

³³ Bjola C., Manor I. First contact: integrating generative AI into digital diplomatic intelligence // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 315–318. – DOI 10.1057/s41254-024-00333-w.

³⁴ The Role of AI in the Battle Against Disinformation [Электронный ресурс] // NATO Strategic Communications Centre of Excellence. URL: <https://stratcomcoe.org/publications/the-role-of-ai-in-the-battle-against-disinformation/238> (дата обращения: 15.04.2026).

ИИ-инструменты или закупать и внедрять готовые аутсорсные решения³⁵. Исследования по цифровой дипломатии показывают, что создание персонализированных алгоритмических решений для нужд дипломатической службы требует не только значительных финансовых вложений на разработку и поддержание системы, но и устойчивых кадровых и организационных ресурсов для сопровождения и обновления этих систем³⁶. При этом такой подход может позволить автоматизировать процессы цифровой дипломатии, сохраняя при этом контроль качества и целенаправленности генерируемого контента³⁷. Для большинства стран, особенно с ограниченными бюджетами, создание собственных ИИ-систем, вероятно, окажется слишком затратным. Появление специальных государственных руководств, таких как NIST AI RMF 1.0³⁸ Министерства торговли США и UK AI Playbook³⁹ правительства Великобритании свидетельствует о сложности и ресурсоемкости институционализации ИИ в государственном секторе.

Одной из основных проблем использования генеративных моделей ИИ является их на данный момент неспособность корректно работать с информацией, если она касается таких чувствительных тем, как войны, гуманитарные кризисы, наркоторговля, продажа оружия, смерти⁴⁰. В таких случаях недостаточно передачи фактов, необходимо использовать уместный тон, учитывать историческую память, культурный контекст⁴¹. Поэтому в

³⁵ Manor I. From digital diplomacy to AI-driven diplomacies? Mapping the potential impact of AI on diplomacy // *Communication and the Public*. – 2026. – Advance online publication. – DOI 10.1177/20570473261438571.; Beulen E., Plugge A., van Hillegersberg J. Formal and relational governance of artificial intelligence outsourcing // *Information Systems and e-Business Management*. – 2022. – Vol. 20. – P. 719–748. – DOI 10.1007/s10257-022-00562-7.

³⁶ Sevin E., Eken M. E. Yet another turn? Prioritising the needs of diplomacy over the capabilities of generative AI // *Place Branding and Public Diplomacy*. – 2025. – Vol. 21. – P. 309–314. – DOI 10.1057/s41254-024-00325-w.

³⁷ «Цифра» и искусственный интеллект на службе дипломатии: аналитический доклад / Е. С. Зиновьева, Н. А. Цветкова, Э. Л. Сидоренко [и др.]; под ред. Е. С. Зиновьевой; Моск. гос. ин-т междунар. отношений (ун-т) МИД России, Каф. мировых полит. процессов. – Москва: МГИМО-Университет, 2024. – 68 с. – ISBN 978-5-9228-2883-3.

³⁸ NIST AI 100-1. Artificial Intelligence Risk Management Framework [Электронный ресурс] / National Institute of Standards and Technology, U.S. Department of Commerce, 2023. 48 с. URL: <https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf> (дата обращения: 15.04.2026).

³⁹ AI Playbook for the UK Government [Электронный ресурс] // UK Government Digital Service, 2025. URL: <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government> (дата обращения: 15.04.2026).

⁴⁰ Ulloa R., Zucker E. M., Bultmann D. [et al.] From prosthetic memory to prosthetic denial: auditing whether large language models are prone to mass atrocity denialism // *AI & Society*. – 2025. – DOI 10.1007/s00146-025-02719-7.

⁴¹ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // *Place Branding and Public Diplomacy*. – 2025. – Vol. 21. – P. 327–333. – DOI 10.1057/s41254-024-00324-x.

вопросах применения ИИ должно четко разделяться, какие темы и задачи ему под силу, а какие нет.

Одной из наиболее востребованных возможностей, открывающихся перед внешнеполитическими ведомствами, выступает расширение доступности консульских услуг и информации для граждан. Традиционно получение консульской поддержки сопряжено с необходимостью личного визита, телефонного звонка или поиска информации на сайте, что создает временные барьеры. Внедрение интеллектуальных диалоговых систем способно трансформировать данный процесс, обеспечивая мгновенный доступ к информации в режиме реального времени⁴².

Зависимость от коммерческих решений в сфере ИИ-технологий создает для внешнеполитических ведомств комплекс рисков. В связи с этим актуализируется возможность создания суверенных систем ИИ, специально адаптированных к потребностям дипломатической практики. Такие системы способны учитывать национально-культурную специфику, языковые особенности, приоритеты внешней политики и требования к защите конфиденциальной информации, что принципиально важно для обеспечения информационного суверенитета государства в цифровой сфере. Практические шаги в данном направлении уже предпринимаются на международном уровне. 1 октября 2025 года Великобритания объявила о международных инициативах «AI for development» на площадке G20, демонстрируя связку дипломатии и ИИ повестки развития⁴³. Проект UNDP по «Digital Economic Diplomacy» для Узбекистана направлен на развитие возможностей МИД в цифровой трансформации внешнеполитического блока⁴⁴. Однако даже

⁴² «Цифра» и искусственный интеллект на службе дипломатии: аналитический доклад / Е. С. Зиновьева, Н. А. Цветкова, Э. Л. Сидоренко [и др.]; под ред. Е. С. Зиновьевой; Моск. гос. ин-т междунар. отношений (ун-т) МИД России, Каф. мировых полит. процессов. – Москва: МГИМО-Университет, 2024. – 68 с. – ISBN 978-5-9228-2883-3.

⁴³ UK announces major AI for development initiatives at G20 in South Africa [Электронный ресурс] // GOV.UK. URL: <https://www.gov.uk/government/news/uk-announces-major-ai-for-development-initiatives-at-g20-in-south-africa> (дата обращения: 15.04.2026).

⁴⁴ Digital Economic Diplomacy [Электронный ресурс]: Enhancing Capacities for Digital Economic Diplomacy in Uzbekistan / United Nations Development Programme (UNDP). URL: <https://www.undp.org/uzbekistan/projects/completed-digital-economic-diplomacy> (дата обращения: 15.04.2026).

создание специализированных систем не устраняет фундаментальных угроз, присущих ИИ-технологиям, прежде всего, феномена «галлюцинаций» и генерации недостоверной информации.

ИИ-«галлюцинации» - одна из главных проблем внедрения ИИ в цифровую дипломатию. Под этим термином подразумевается способность генеративных моделей создавать тексты, изображения и прочие материалы, которые кажутся правдоподобными, логично построенными и даже стилистически выверенными, но на деле содержат ошибки, вымысел или искажение реальности⁴⁵. В цифровой дипломатии доверие к официальным источникам информации и точность формулировок крайне важны, но эти качества могут быть подорваны⁴⁶. Практическое значение проблемы подтверждается специализированными актами, например: меморандум Офиса по вопросам управления и бюджету при президенте США от 28 марта 2024 года предписывает обязательное ограничение или остановку ИИ-кейсов в правительстве США в случае обнаружения риска для безопасности⁴⁷.

Не менее значим риск нарушения конфиденциальности служебной информации. Публичные ИИ-сервисы функционируют на основе облачной инфраструктуры, при этом данные, загружаемые пользователем, становятся доступными компании разработчику и могут использоваться для обучения модели. Это создает угрозу утечки конфиденциальной информации. Наглядным примером служит инцидент, произошедший в июле 2025 года, когда исполняющий обязанности директора Агентства по кибербезопасности и защите инфраструктуры США (CISA) Мадху Готтумуккала загрузил в ChatGPT служебные документы помеченные грифом «For Official Use Only»⁴⁸. Подобного рода инциденты служат подтверждением того, что угроза

⁴⁵ Галлюцинация искусственного интеллекта [Электронный ресурс] // Лаборатория Касперского. URL: <https://www.kaspersky.ru/resource-center/definitions/ai-hallucination> (дата обращения: 15.04.2026).

⁴⁶ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 327–333. – DOI 10.1057/s41254-024-00324-x.

⁴⁷ M-24-10. Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence [Электронный ресурс] / Office of Management and Budget, Executive Office of the President. – 2024. – 29 с. – URL: <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf> (дата обращения: 15.04.2026).

⁴⁸ Trump's acting cyber chief uploaded sensitive files into a public version of ChatGPT [Электронный ресурс] // POLITICO. 2026. 27 Jan. URL: <https://www.politico.com/news/2026/01/27/cisa-madhu-gottumukkala->

нарушения конфиденциальности данных при использовании ИИ-технологий носит не гипотетический, а реальный характер.

Еще одной угрозой внедрения ИИ-систем выступает феномен «черного ящика». Под данным понятием понимается непрозрачность внутренних механизмов принятия решений нейросетевыми моделями. Исследователи и пользователи имеют доступ к входным данным и выходным результатам, однако процессы трансформации одних в другие остаются недоступны для непосредственного наблюдения и интерпретации⁴⁹. Это обстоятельство создает угрозу для дипломатических структур, традиционно предполагающих необходимость обоснования и документирования принимаемых решений. Когда ИИ генерирует аналитическую справку, проект заявления или рекомендацию по тому или иному вопросу, ИИ не всегда способна объяснить, на основании каких именно данных и логических операций был получен соответствующий результат, что подрывает легитимность использования таких материалов в официальной коммуникации⁵⁰.

Таким образом, интеграция ИИ в практику цифровой дипломатии характеризуется выраженной двойственностью: технологии открывают новые возможности для повышения эффективности внешнеполитической деятельности, однако одновременно порождают комплекс рисков, способных нивелировать достигнутые преимущества.

Проверка постов внешнеполитических ведомств системами детекции сгенерированного текста.

В целях проверки текущего использования ИИ для генерации постов была проведена пилотная проверка текстов Telegram каналов внешнеполитических ведомств сервисами обнаружения ИИ-сгенерированных текстов. Были выбраны 2 канала - *Russian MFA*, официальный канал МИД РФ на английском языке и *США по-русски*, официальный русскоязычный аккаунт

[chatgpt-00749361](#) (дата обращения: 15.04.2026).

⁴⁹ Pisoni G., Moloney M. Responsible AI-Based Business Process Management and Improvement // DISO. – 2024. – Vol. 3. № 23. – DOI 10.1007/s44206-024-00105-2.

⁵⁰ Huang Z. A. Terminology, AI bias, and the risks of current digital public diplomacy practices // Place Branding and Public Diplomacy. – 2025. – Vol. 21. – P. 327–333. – DOI 10.1057/s41254-024-00324-x.

Государственного Департамента США. Публикации выбирались методом случайного отбора с помощью генератора случайных чисел на сайте RANDOM.ORG⁵¹, за период январь - февраль 2026 года. Каждую неделю выбирался день недели, затем сам пост, любой графический контент, публикации содержащие цитаты или пересланные сообщения в выборку не включались. Если в выбранный день отсутствовали публикации, алгоритм поиска начинался заново, исключая день, в котором не была обнаружена подходящая публикация. Итоговый массив данных составил 18 публикаций, по 9 от каждого канала.

Для оценки степени участия ИИ в генерации текста было выбрано 5 детекторов ИИ - ISGEN⁵², Domate AI⁵³, QuillBot⁵⁴, Originality AI⁵⁵ и GPTZero⁵⁶. Выбор именно этих систем проверки текстов на наличие сгенерированных текстов обусловлен их единой функциональной направленностью и методами проверки, включающими лингвистический, статистический и вероятностный анализ. Кроме того, для решения поставленной задачи эти системы позволяли проверять тексты без оплаты и дополнительных настроек доступа.

Результаты проверки представлены в Таблице 2, данные отражают средние значения процента сгенерированного текста при помощи ИИ-технологий. В связи с тем, что и опыт проверок, и утверждения разработчиков сервисов детекции ИИ о том, что результаты детекции сильно зависят от языка (детекция хуже на неанглийских текстах), длины фрагментов и других параметров, мы понимаем, что в результате могут быть как ложноотрицательные, так и ложноположительные срабатывания. Тем не

⁵¹ RANDOM.ORG – сервис генерации случайных чисел на основе атмосферного шума [Электронный ресурс]. URL: <https://www.random.org> (дата обращения: 15.04.2026).

⁵² ISGEN.AI – многоязычный детектор контента, сгенерированного ИИ британской компании ISGEN LTD [Электронный ресурс]. URL: <https://isgen.ai/ru> (дата обращения: 15.04.2026).

⁵³ Domate AI – система анализа текста на заимствования и ИИ-контент российской компании Domate LLC [Электронный ресурс]. URL: <https://app.domate.ru> (дата обращения: 15.04.2026).

⁵⁴ QuillBot – платформа, сочетающая множество инструментов, включая ИИ-детектор американской компании QuillBot [Электронный ресурс]. URL: <https://quillbot.com> (дата обращения: 15.04.2026).

⁵⁵ Originality.ai – детектор контента, сгенерированного ИИ и проверка на плагиат американской компании Originality.ai [Электронный ресурс]. URL: <https://originality.ai> (дата обращения: 15.04.2026).

⁵⁶ GPTZero – детектор контента, сгенерированного ИИ американской компании GPTZero [Электронный ресурс]. URL: <https://gptzero.me> (дата обращения: 15.04.2026).

менее, используя одни и те же сервисы проверки, мы получили интересные результаты.

Таблица 2. Результаты проверки постов на наличие текста, сгенерированного ИИ.

Детектор ИИ	Russian MFA	США по-русски
ISGEN	15%	25%
Domate AI	0%	14%
QuillBot	13%	0%
Originality AI	9%	1%
GPTZero	0%	0%

Результаты проверки показали, что уровень детекции признаков ИИ-генерации варьируются в диапазоне от 0% до 25% в зависимости от используемого детектора. Зафиксированные результаты не позволяют сделать однозначный вывод о генерации текста при помощи технологий ИИ, в большинстве классификаций сервисов верификаторов значения до 25% соответствуют низкой вероятности машинного генерации текста.

Таким образом, для анализируемых каналов цифровой дипломатии вероятно, остается характерным преимущественно ручное, традиционное создание текста, при котором ИИ-инструменты если и используются, то носят исключительно вспомогательный характер.

Интерпретация результатов через призму этических принципов Рекомендации ЮНЕСКО об этике использования ИИ.

Полученные результаты целесообразно интерпретировать в рамках матрицы Рекомендаций ЮНЕСКО об этике искусственного интеллекта 2021 года⁵⁷. Принцип человеческого достоинства показывает ограниченность коммерческих ИИ моделей в работе с чувствительными темами, когда защитные ограничения могут снижать вред, а в кризисных ситуациях

⁵⁷ Recommendation on the Ethics of Artificial Intelligence [Электронный ресурс] / UNESCO. – 2022. – 43 p. – URL: <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics> (дата обращения: 15.04.2026).

создавать информационный вакуум. Принцип пропорциональности требует, чтобы в вопросах высокой общественной значимости итоговое решение принимал человек.

В пунктах 37-41 содержатся принципы прозрачности, которые делают проблемой непрозрачность алгоритмов, на основе которых работают ИИ и отсутствие маркировки ИИ-контента. Феномен «ИИ-галлюцинаций» рассматривается через призму ответственности в пунктах 42-43. Всегда в конечном счете ответственность должна возлагаться на участников ИИ-процесса в соответствии с их ролью в жизненном цикле системы. В дипломатической практике, где доверие к официальным источникам информации является базовой необходимостью, принцип играет особую роль. Государство несет полную ответственность за любой контент, опубликованный от его имени, независимо от того, был он создан человеком или сгенерирован алгоритмом. Невозможность гарантировать достоверность ИИ-материалов ставит под угрозу легитимность официальной коммуникации.

Результаты пилотной проверки постов в Telegram-каналах МИД РФ и Госдепартамента США показали преимущественно ручное создание текстовых публикаций. Данный факт может быть рассмотрен как подтверждение принципа человеческого контроля, закрепленного в пунктах 35-36. Согласно рекомендации, государства должны обеспечить возможность определить этическую или юридическую ответственность за любой этап жизненного цикла ИИ. При этом, на сегодняшний день маловероятно, что ИИ сможет заменить окончательный человеческий контроль и ответственность.

Рекомендации об этике искусственного интеллекта представляют собой наиболее авторитетную на сегодняшний день международную рамку, применимую к оценке внедрения ИИ в различные сферы жизни человека. Ее ценность заключается в том, что документ не предлагает жестких запретов, а формулирует систему ценностей и принципов. Документ опирается на нормы международного права, включая международные пакты о правах человека, что придает положениям универсальный характер и легитимность.

Заключение.

Таким образом, интеграция технологий ИИ в сферу цифровой дипломатии носит двойственный характер. С одной стороны, ИИ открывает значительные возможности для повышения эффективности внешнеполитической деятельности, с другой порождает серьезные риски, способные подрвать доверие к официальным источникам информации и информационному суверенитету государства. Сфера цифровой дипломатии, в отличие от многих других сфер общественной и политической жизни, где внедрение ИИ становится стандартом, является более уязвимой, т.к. последствия рисков пока сложно просчитать.

Тем не менее, сегодня уже можно сформулировать рекомендации, которые могут нивелировать возможные риски внедрения ИИ в цифровую дипломатию:

1. Между ИИ и человеком необходимо четко делегировать задачи. Рутинные технические операции делегировать ИИ, а все вопросы, связанные с чувствительными темами, подготовкой заявлений или иного контента, должны оставаться исключительно в компетенции дипломатического персонала.
2. Разрабатывать и внедрять суверенные или национально адаптированные ИИ-системы, которые будут учитывать специфику внешней политики, языковые и культурные особенности. Для государств с ограниченными ресурсами рекомендуется либо вовсе отказаться от использования открытых международных ИИ-инструментов, либо использовать ИИ-системы «дружественных стран».
3. При использовании любых систем ИИ необходимо запретить или максимально ограничить загрузку служебной или конфиденциальной информации. Требуется разработать и внедрить внутренние политики использования ИИ для дипломатического персонала.

4. Проводить регулярное обучение дипломатов использованию ИИ-технологий, включая мировые практики, этические принципы и специфику деятельности.

Да, ИИ способен существенно усилить инструментарий цифровой дипломатии, повысив ее оперативность, охват и ресурсную эффективность. Однако это возможно только при условии строгого соблюдения принципа человеческого контроля, прозрачности и ответственности. Последовательная реализация предложенных рекомендаций позволит грамотно реализовать преимущества новых технологий, минимизируя при этом риски для национальной безопасности, доверия к официальным источникам и легитимности дипломатической коммуникации.

В дальнейших исследованиях этой темы видится проведение полноценного контент-анализа материалов, содержащихся в официальных Telegram-каналах внешнеполитических ведомств, с целью выявления устойчивых, регулярных практик использования ИИ в цифровой дипломатии и перейти от постановки проблемы к доказательному исследованию роли генеративных моделей ИИ в дипломатическом дискурсе.