

УДК 070.1:004.738.5

DOI: 10.26456/vtfilol/2024.3.172

НОВЫЕ ВЫЗОВЫ МЕДИЙНОГО ФАКТЧЕКИНГА НА ФОНЕ РОСТА ГЕНЕРАТИВНОГО ИИ

А. В. Соколов

Академия медиаиндустрии, г. Москва

Статья посвящена проблеме распространения дезинформации в медиа и социальных сетях, где искусственный интеллект играет все более значимую роль. Автор анализирует эффективность фактчекинга как средства борьбы с фальшивыми новостями и дезинформацией в медиа, а также обсуждают перспективы развития технологий для оценки достоверности информации.

Ключевые слова: ИИ, дезинформация, фактчекинг, медиа.

Введение, постановка проблемы

Цифровизация современного медийного пространства способствуют распространению дезинформации, а массовый доступ к ИИ начинает играть все более значительную роль в этом процессе [20]. В то же время с ростом масштаба дезинформации в медиа и социальных сетях ИИ-решения для проверки информации рассматриваются как возможный «святой Грааль» фактчекинга [14].

В 2024 году в более чем 50 странах [22] мира проходят выборы, что привлекает внимание специалистов к кампаниям по дезинформации, использующим поддельные изображения, фальшивые видео и аудиозаписи, фейковые новости, в том числе созданные генеративным искусственным интеллектом [13]. Легкий доступ к инструментам искусственного интеллекта рискует вызвать «приливную волну политически мотивированной лжи, которая заполнит каналы социальных сетей способами, которые казались невообразимыми даже несколько лет назад» [24]. Большинство пользователей не способны уверенно отличить контент, сгенерированный ИИ [2; 27; 15; 19]. Исследование, опубликованное в журнале *Science Advances*, показывает, что ИИ может быть эффективней людей в создании дезинформации [26]. Дж. Спитале из Цюрихского университета, руководивший исследованием, отмечает: «Тот факт, что дезинформация, генерируемая ИИ, не только дешевле и быстрее, но и более эффективна, вызывает у меня кошмары» [29].

© Соколов А. В., 2024

Фактчекинг: предпосылки и анализ актуальных исследований

В некоторых работах, где прямо рассматриваются этические аспекты фальшивых новостей, они концептуализируются как дезинформация [25, с. 180]. Дезинформацию можно определить как «намеренное создание и распространение ложной и/или вводящей в заблуждение информации» [10, с. 1]. В противоположность этому, фактчекинг направлен на верификацию и коррекцию недостоверной информации, обеспечивая доступ к проверенным и точным данным.

Проверка фактов всегда была частью журналистского процесса (так называемая «внутренняя» или *ante hoc* проверка), но как отдельные новостные материалы («внешняя» или *post hoc* проверка) она начала развиваться только с 2000-х годов [28] и теперь включает в себя «функции выявления и развенчания и относится к широкому кругу публикаций/информации, особенно к контенту социальных сетей» [18]. Некоторые авторы исследуют *post hoc* фактчекинг как новый жанр журналистики [28].

Современные подходы к внешней проверке фактов возникли с развитием Интернета и как реакция на президентские выборы в США в 1988 году. К.Х. Джеймисон [12] вместе с Б. Джексоном создала *FactCheck.org* в 2003 году. В 2007 году появилась *PolitiFact.com* и *Fact Checker* от *The Washington Post*, став ведущими политическими фактчекерами в США [1]. В последующие годы было создано несколько организаций по проверке фактов, например, *Full Fact*, *Snopes*, *Poynter* и *NewsGuard*. Таким образом, проверка фактов изначально шла рука об руку с освещением выборов и политической журналистикой. Однако с развитием социальных сетей все больше внимания стало уделяется разоблачению онлайн-дезинформации, вирусных мемов и другого мистификационного контента [11]. Это стало актуальным в связи с референдумом по *Brexit* в 2016 году и президентскими выборами в США в том же году. В 2020 и 2021 годах эта проблема вновь оказалась в центре внимания в связи с беспрецедентным распространением дезинформации, связанным с *COVID-19* и результатами президентской гонки в США 2020 года [7]. Представляется, что «самая главная опасность, связанная с фальшивыми новостями, заключается в том, что они обесценивают и делегитимизируют голоса экспертов, авторитетные институты и концепцию объективных данных» [Там же]. Все это вызвало девятикратный рост в период с 2010 по 2021 год количества инициатив по проверке фактов, на 2021 год в 106 странах мира насчитывалось 341 команда фактчекеров [Там же].

Проблемы дезинформации в медиа и фактчекинга на протяжении последнего десятилетия находится в фокусе внимания как ученых, так и практиков медиаиндустрии. Так, с 2014 проходят ежегодные саммиты *GlobalFact*, организованные Международной сетью проверки фактов (*IFCN*) при Институте Пойнтера по всему миру, 11-й саммит состоится в столице Боснии и Герцеговины Сараево в июне 2024 года [9]. С 2019 года онлайн-конференция «Правда и доверие в интернете» (*Truth and Trust Online*) проходит в разных странах, объединяя ученых, представителей промышленности, некоммерческие организации [6]. В 2022 году в Афинах (Греция) прошла первая международная научная конференция по фактчекингу (*Tackling Misinformation: The Contribution of Universities and Research Centers*), посвященная вкладу университетов и исследовательских центров в борьбу с дезинформацией [30]. В октябре 2022 года Гонконгский баптистский университет организовал виртуальную конференцию «Проверка фактов: глобальная перспектива» (*Checking the Fact Checkers: A Global Perspective*) [23]. В 2023 году состоялся Кембриджский саммит по дезинформации в *King's College (Cambridge Disinformation Summit)* [3]. В феврале 2024 года в университете Амстердама прошла научная конференция *EDMO Scientific Conference 2024*, посвященная проблемам дезинформации [8].

В последнее время исследователи фокусируются на разработке ИИ-технологий для оценки достоверности источников информации, в том числе при финансовой поддержке международных фондов [21; 31].

Разрабатываются методы, основанные на нейронных сетях, для анализа текстовых статей, учитывающих семантическое сходство, тональность и надежность источника [4]. Другое направление – вычислительная проверка фактов, которая может значительно расширить возможности оценки достоверности информации [5]. Также разрабатываются методы автоматического обнаружения ложных новостей в интернете и социальных сетях [16]. В целом, речь идет о подходах, основанных на автоматическом обнаружении фейковых новостей и автоматизированной и вычислительной проверке фактов.

Выводы

В целом на современном этапе можно говорить об использовании разнообразного набора инструментов проверки фактов. Автоматизированные инструменты предоставляют скорость и масштабируемость, но могут страдать от неточностей и алгоритмических предвзятостей. Краудсорсинговые платформы вовлекают сообщества в процесс проверки, хотя и имеют проблемы с координацией и надежностью. Профессиональные

организации по проверке фактов обеспечивают экспертность и тщательность, но их возможности ограничены ресурсами. Образовательные программы по медиаграмотности способствуют долгосрочному повышению критического мышления у населения. Гибридные модели, объединяющие автоматизированные инструменты, краудсорсинг и профессиональный надзор, предлагают наиболее всесторонний подход, хотя и требуют сложной координации. Эти подходы различаются по принципу используемых ресурсов и методов, а также по вовлечению и роли различных субъектов (технологии, сообщество, эксперты, образование) [17] в процессе фактчекинга.

Список литературы

1. Amazeen M. The promise and pitfalls of fact-checking in 2022 // *Science and Society: Journal of Political and Moral Theory*. 2024. Vol. 43. P. 11–33.
2. Barthel M., Mitchell A., Holcomb J. Many Americans believe fake news is sowing confusion // *Pew*: [website]. 2016. 15 December. URL: <http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/> (accessed: 25.05.2024).
3. Cambridge Disinformation Summit – CFRA conferences and events // *University of Cambridge Judge Business School*: [website]. URL: <https://www.jbs.cam.ac.uk/faculty-research/centres/cfra/conferences-events/cambridge-disinformation-summit/> (accessed: 02.06.2024).
4. Choudhary N., Singh R., Bindlish I., Shrivastava M. Neural network architecture for credibility assessment of textual claims // *ArXiv:1803.10547*. URL: <http://arxiv.org/abs/1803.10547> (accessed: 22.05.2024).
5. Ciampaglia GL, Shiralkar P, Rocha LM, Bollen J, Menczer F, Flammini A (2015) Computational Fact Checking from Knowledge Networks // *PLoS ONE*. No. 10(6): e0128193. Doi:10.1371/journal.pone.0128193.
6. Conference for Truth and Trust Online – 2022. URL: <https://truthandtrustonline.com/> (accessed: 23.05.2024).
7. Damme T. V. Global Trends in Fact-Checking: A Data-Driven Analysis of Claim-Review // *Master's Thesis (University of Antwerp)*. 2021. URL: https://www.academia.edu/59259728/Global_Trends_in_Fact_Checking_A_Data_Driven_Analysis_of_ClaimReview?rhid=28619727326&swp=rr-rw-wc-114394632 (accessed: 02.06.2024).
8. EDMO Scientific Conference 2024 // *EDMO*: website. 2024. URL: <https://edmo.eu/event/edmo-scientific-conference-2024/> (accessed: 22.04.2023).
9. GlobalFact 11 // *Poynter*: website. 2024. URL: <https://www.poynter.org/event/globalfact-11/> (accessed: 22.05.2024).
10. Hameleers M. Disinformation as a context-bound phenomenon: toward a Conceptual Clarification Integrating actors, Intentions and Techniques of Creation and Dissemination // *Communication Theory*. 2022. T. 33. № 1. Pp. 1–10.

11. Holan A. D. Fact-checkers and the social media misinformation tsunami: A Q&A with Lucas Graves // Poynter : website. 2023. URL: <https://www.poynter.org/fact-checking/2023/fact-checking-online-misinformation-vs-politics-government/> (accessed: 22.05.2024).
12. Jamieson K. H. Dirty Politics: Deception, Distraction, and Democracy. Oxford University Press, 1993. 335 p.
13. Kahn G. Generative AI Is Already Helping Fact-Checkers. But It's Proving Less Useful in Small Languages and Outside the West // Reuters Institute for the Study of Journalism : website. 2024. 29 Apr. URL: <https://reutersinstitute.politics.ox.ac.uk/news/generative-ai-already-helping-fact-checkers-its-proving-less-useful-small-languages-and> (accessed: 27.05.2024).
14. Kavtaradze L., Kalsnes B. AI-powered fact-checking: Strategic framing of AI use for information verification // Strategic communication – contemporary perspectives / Eds. S. Alghasi, E. C. Vanvik, J. Barland, J. Falkheimer. 2024. Ch. 9. Pp. 177–198. <https://doi.org/10.23865/noasp.208.ch9>.
15. Kumar R., Mindzak M. Who Wrote This? Detecting Artificial Intelligence–Generated Text from Human-Written Text // Canadian Perspectives on Academic Integrity. 2024. № 7(1). <https://doi.org/10.55016/ojs/cpai.v7i1.77675>.
16. Lahlou Y., El Fkihi S., Faizi R. Automatic detection of fake news on online platforms: A survey, 2019 1st International Conference on Smart Systems and Data Science (ICSSD), Rabat, Morocco, 2019. Pp. 1-4. Doi: 10.1109/ICSSD47982.2019.9002823.
17. Monnier A. A Typology of Fact-Checking Resources. From False/True Verification to Information Literacy // Questions de communication. 2023. Vol. 43. No. 1. Pp. 297-316.
18. Monnier A. Mapping Fact-Checking Resources. A Typology Based On Cross-National Insights (France-Greece-Poland) // Findings Report – 2022 CALYPSO. 2022. URL: https://edmo.eu/wp-content/uploads/2022/07/Calypso_report_v3.pdf (accessed: 24.05.2024).
19. Naaman M. «Silly» to expect AI mimicking human-written text could be easily detected // Cornell Chronicle : website. URL: <https://news.cornell.edu/media-relations/tip-sheets/silly-expect-ai-mimicking-human-written-text-could-be-easily-detected> (accessed: 28.05.2024).
20. Navigating the AI-Generated Information Landscape: Finding the Truth amid Misinformation // LexisNexis : website. 2023. 27 Febr. URL: www.lexisnexis.com/community/insights/professional/b/industry-insights/posts/misinformation-in-artificial-intelligence- (accessed: 26.05.2024).
21. Network T. I. F.-C. International Fact-Checking Network awards \$2 million to 20 groups for major fact-checking projects // Pointer : website. 2024. URL: <https://www.poynter.org/fact-checking/2024/international-fact-checking-network-awards-2-million-to-20-groups-for-major-fact-checking-projects/> (accessed: 28.05.2024).

22. Ray S. 2024 Is the Biggest Election Year in History – Here Are the Countries Going to the Polls This Year // *Forbes* : website. 2024. URL: www.forbes.com/sites/siladityaray/2024/01/03/2024-is-the-biggest-election-year-in-history-here-are-the-countries-going-to-the-polls-this-year/ (accessed: 22.05.2024).
23. School of Communication, Hong Kong Baptist University – Checking the Fact Checkers: A Global Perspective // Hong Kong Baptist University : website. 2022. URL: <https://www.comm.hkbu.edu.hk/aimr/checking-the-fact-checkers/index.php> (accessed: 22.05.2024).
24. Scott M. Deepfakes, distrust and disinformation: Welcome to the AI election // *Politico* : website. 2024. URL: <https://www.politico.eu/article/deepfakes-distrust-disinformation-welcome-ai-election-2024/> (accessed: 22.04.2023).
25. Scott R. Stroud Pragmatist Media Ethics and the Challenges of Fake News // *Journal of Media Ethics*. 2019. Vol. 34. No. 4. Pp. 178–192.
26. Spitale G., Biller-Andorno N., Germani F. AI model GPT-3 (dis)informs us better than humans // *Science Advances*. 2023. T. 9. № 26. eadh1850. Doi: 10.1126/sciadv.adh1850.
27. Study Finds Majority of People Unable to Distinguish AI-Generated Text from Human Writing // *AI-PRO.org* : website. 2023. URL: <https://ai-pro.org/study-finds-majority-of-people-unable-to-distinguish-ai-generated-text-from-human-writing/> (accessed: 28.05.2024).
28. Van Der Meulen M., Reijnierse W. European Language Resources Association (ELRA), licensed under CC., 2020. URL: <https://repository.uibn.ru.nl/bitstream/handle/2066/219847/219847.pdf?sequence=1&isAllowed=y> (accessed: 02.06.2024).
29. Williams R. Humans May Be More Likely to Believe Disinformation Generated by AI // *Technology Review* : website. 2023. URL: <https://www.technologyreview.com/2023/06/28/1075683/humans-may-be-more-likely-to-believe-disinformation-generated-by-ai/> (accessed: 27.05.2024).
30. 1st International Scientific one-day conference on fact-checking in Greece Tackling Misinformation: The Contribution of Universities and Research Centers. URL: <https://diavlos.grnet.gr/en/event/e4039> (accessed: 27.05.2024).
31. 8 Projects Approved Under Fact-Checking’s Sixth Funding-Round. URL: <https://gulbenkian.pt/emifund/news/8-projects-approved-under-fact-checkings-sixth-funding-round/> (accessed: 27.05.2024).

EMERGING CHALLENGES IN MEDIA FACT-CHECKING AMIDST THE RISE OF GENERATIVE AI

A. V. Sokolov

Academy of Media Industry, Moscow

The article is devoted to the problem of misinformation dissemination in media and social networks, where artificial intelligence is playing an increasingly important role. The author analyzes the effectiveness of fact-checking as a means

of combating fake news and misinformation, and discusses the prospects for the development of technologies for evaluating the reliability of information.

Keywords: *AI, disinformation, fact-checking, media.*

Об авторе:

СОКОЛОВ Александр Владимирович – кандидат филологических наук, профессор кафедры телевидения и радиовещания Академии медиаиндустрии (127521, Москва, ул. Октябрьская, д. 105, корп. 2), e-mail: sokolov.ipk@internet.ru.

About the author:

SOKOLOV Aleksandr Vladimirovich – Candidate of Philology, Professor at the Department of Television and Radio Broadcasting, Academy of Media Industry (127521, Moscow, October str., 105, k. 2), e-mail: sokolov.ipk@internet.ru.

Дата поступления рукописи в редакцию: 01.09.2024 г.

Дата подписания в печать: 06.09.2024 г.