

4. Федеральный закон от 31 июля 2020 г. № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации» // СПС «КонсультантПлюс» [Электронный ресурс]. – URL: https://www.consultant.ru/document/cons_doc_LAW_358738/
5. Распоряжение Правительства Российской Федерации от 19 августа 2020 г. № 2129-р «Об утверждении Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года» // СПС «КонсультантПлюс» [Электронный ресурс]. – URL: https://www.consultant.ru/document/cons_doc_LAW_360681/
6. Кутейников Д. Л., Ижаев О. А., Зенин С. С., Лебедев В. А. Ключевые подходы к правовому регулированию использования систем искусственного интеллекта // Вестник Тюменского государственного университета. Социальноэкономические и правовые исследования. – 2022. – Т.8. – № 1 (29). – С. 209–232.
7. DARPA и наука Третьего рейха : оборонные исследования США и Германии / Под общ. ред. А.Е. Суворова. – М.: ТЕХНОСФЕРА, 2020. – С. 122– 123.
8. Егорова Н. Е., Торжевский К. А. Методологические основы принятия стратегических решений. Секция 1. «Теоретические проблемы стратегического планирования на микроэкономическом уровне». Стратегическое планирование и развитие предприятий: материалы XXIII Всероссийского симпозиума. Москва, 12–13 апреля 2022 г. / под ред. чл.-корр. РАН Г. Б. Клейнера. М.: ЦЭМИ РАН, 2022. – С. 25–28.
9. Варшавский Л. Е. Социально-экономические проблемы развития информационно-коммуникационных технологий (ИКТ): монография. – М.: ЦЭМИ РАН, 2022.
10. Самбунова Т. Методологические вопросы анализа цифровой экономики [Электронный ресурс] // URL: <https://delyagin.ru/articles/191-materialy-mgd/110905-metodologicheskie-voprosy-analiza-tsifrovoy-jekonomiki>

И. Г. Ильин,

аспирант,

Санкт-Петербургский государственный университет

СОЗДАНИЕ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ И ИХ ВЛИЯНИЕ НА РЕАЛИЗАЦИЮ И ЗАЩИТУ КОНСТИТУЦИОННО ГАРАНТИРОВАННЫХ ПРАВ

Аннотация. Статья посвящена исследованию некоторых правовых аспектов создания больших языковых моделей в контексте цифрового неравенства и языковой дискриминации. В результате анализа было установлено, что применительно к большим языковым моделям цифровое неравенство возникает из-за неравномерного уровня цифровизации языков и проявляется в ограниченном доступе к технологии обработки естественного языка.

Ключевые слова: право, большие языковые модели, языковая дискриминация, цифровые технологии, цифровое неравенство, искусственный интеллект, технология обработки естественного языка

CREATION OF LARGE LANGUAGE MODELS AND THEIR IMPACT ON THE REALIZATION AND PROTECTION OF CONSTITUTIONALLY GUARANTEED RIGHTS

Abstract. The lecture notes focus on the results of research examining the constitutional and legal aspects of large language model development in the context of the digital divide and linguistic discrimination. The study concludes that, in relation to the development of large language models, the digital divide arises due to the uneven digitalization of languages, resulting in limited access to natural language processing technology.

Keywords: law, large language models, language discrimination, digital technologies, digital divide, artificial intelligence, natural language processing technology

Большие языковые модели (англ. Large Language Models, LLM) – это генеративные модели искусственного интеллекта используемые в технологии обработки естественного языка (англ. Natural language processing, NLP). Наличие таких моделей позволяет компьютеру эффективно обрабатывать текстовые данные, демонстрируя способность к «пониманию» текста на глубоком уровне, создавать связные и контекстуально релевантные ответы на запросы, осуществлять перевод текста между языками, а также генерировать текст, который соответствует определенным стилевым и содержательным требованиям [18. С. 24–34].

Большие языковые модели обучаются на обширных массивах языковых данных. Структура и репрезентативность, их объем и формат определяют процесс обучения и точность понимания контекста [21. С. 36–56], а наличие дефектов или недостаточность данных может приводить к некорректной работе модели и в целом препятствовать развитию технологии [20. С. 260]. Параметры обучающих данных будут зависеть от уровня цифровизации конкретного языка.

Цифровизация языка в широком смысле представляет собой процесс преобразования данных в соответствующие лингвистические корпуса. Для этого применяются текстовые данные (такие как файлы, транскрипции, аннотации), речевые данные (например, аудиозаписи, фонетические и интонационные аннотации) и мультимодальные данные, которые объединяют несколько типов информации, например видео с текстом, изображения с текстом и т. д. [12. С. 291]. Процесс цифровизации языка выполняет не только техническую функцию, способствуя созданию больших языковых моделей и развитию технологий обработки естественного языка, но и играет важную общественно-социальную роль, содействуя сохранению национальной и культурной идентичности [28. С. 126–139]. Например, цифровизация миноритарных языков может значительно помочь в сохранении культурного наследия малых народов.

Разные языки имеют различный уровень цифровизации, что может быть обусловлено рядом факторов, включая экономические, технические и правовые. Экономические факторы связаны с тем, что языки имеют разный экономический потенциал [6. С. 173; 32. С. 1–8], а процесс цифровизации требует значительных ресурсов, в том числе временных, финансовых и т. д. К техническим факторам может отнести ошибки при сборе данных, недостатки в конструкции корпусов и

ограничения существующих наборов данных, ошибки в метаданных и т. д. [39. С. 193–204; 14. С. 5751–5767; 30. С. 26–37]. Правовые факторы связаны с существованием нормативных барьеров на доступ к обучающим данным, связанных с необходимостью соблюдения соответствующего правового режима при использовании данных [21. С. 36–56; 1. С. 123–140; 22. С. 99–123; 23. С. 367–376; 24. С. 17–27; 25. С. 2–26]. Преодоление таких барьеров требует мер как на концептуальном уровне – устранения нормативных ограничений с учетом баланса частных и публичных интересов, так и на практическом – создания условий для обмена и распространения языковых данных. Среди таких мер можно выделить развитие института повторного использования данных, накопленных в государственных информационных системах, а также привлечение высших учебных заведений к созданию лингвистических корпусов и цифровизации языка.

Невозможность создания полноценной языковой модели для определенных языков или диалектов делает многие цифровые продукты недоступными для их носителей либо ухудшает их работу по сравнению с теми, что доступны для носителей языков с высоким уровнем цифровизации. Это приводит к цифровому неравенству, когда доступ к технологии распределяется неравномерно среди языковых сообществ, тем самым создавая риск дискриминации.

Национальная стратегия развития искусственного интеллекта на период до 2030 года подчеркивает, что защита прав и свобод человека является одним из основных принципов развития и использования технологии искусственного интеллекта, а «недискриминация» выделена в качестве одного из основного принципа развития нормативно-правового регулирования общественных отношений, связанных с развитием и использованием технологий искусственного интеллекта.

Статья 2 Всеобщей декларации прав человека (1948) устанавливает запрет на дискриминацию, в том числе, по языковому признаку. Аналогичное положение содержится и в статье 1 (3) Устава ООН, а также находит свое отражение в пункте 2 статьи 19 Конституции РФ, согласно которому государство гарантирует равенство прав и свобод человека и гражданина независимо от языка. В контексте создания больших языковых моделей представляется целесообразным толковать это понятие максимально широко, рассматривая языковую дискриминацию не только как необоснованные различия или ограничения, которые могут снижать или исключать возможность реализации прав на основе языковой принадлежности, но и как действия, которые могут препятствовать сохранению или развитию языков меньшинств.

Комитет ООН по правам человека неоднократно рассматривал проблему языковой дискриминации, однако его судебная практика недостаточно развита и не всегда обеспечивает надежной защиты языковых меньшинств [31. С. 27–56]. Например, дискриминация по признаку голоса часто остается незамеченной [7. С. 167–177], что может оказаться критичным при взаимодействии с технологией распознавания речи и голоса и связанных цифровых продуктов: систем интерактивного ответа и голосовых помощников. Анализ существующей практики позволяет выделить ряд ключевых критериев для определения языковой дискриминации. Во-первых, это численность носителей языка, поскольку степень дискриминации часто зависит от уровня распространенности языка в обществе

[29. С. 3–15]. Во-вторых, значимым критерием является уровень государственной поддержки многоязычия: активные государственные программы по обучению и использованию различных языков способствуют снижению риска языковой дискриминации [36. С. 39–53]. В-третьих, важен уровень использования языков меньшинств в общественной жизни, что также способствует более точной оценке случаев языковой дискриминации.

Однако, несмотря на возможность выделить критерии для оценки языковой дискриминации, правовая квалификация таких случаев в контексте цифровых технологий вызывает определенные трудности. Например, дискриминация в языковых моделях может быть вызвана как алгоритмической, так и человеческой предвзятостью. Алгоритмическая предвзятость возникает из-за искажений в данных, на которых обучается модель, в то время как человеческая предвзятость может проявляться на этапе разработки алгоритмов [5. С. 337–358]. Оба типа предвзятости могут существенно влиять на точность и справедливость принимаемых решений, что усугубляет социальное неравенство. Важно отметить, что данные формы предвзятости не всегда очевидны и могут проявляться на различных стадиях работы модели, что затрудняет их обнаружение и минимизацию. В связи с этим правовая квалификация таких ситуаций требует особого внимания, поскольку не всегда ясно, следует ли расценивать ошибки в языковых моделях как проявление дискриминации, и какие меры ответственности могут быть применимы к разработчикам и владельцам технологий.

Список литературы

1. Ильин И. Г. Персональные данные в системах искусственного интеллекта: технология обработки естественного языка // *Journal of Digital Technologies and Law*. – 2024. – Т. 2, № 1. – С. 123–140.
2. Миронова М. В. Становление термина «языковая дискриминация» в современной социолингвистике // *New Language*. – New World. – New Thinking. – 2019. С. 555–558.
3. Мушаков В. Е. Конституционные права человека в контексте проблемы преодоления цифрового разрыва // *Вестник Санкт-Петербургского университета МВД России*. – 2022. – № 1(93). – С. 69–73.
4. Талапина Э. В. Обработка данных при помощи искусственного интеллекта и риски дискриминации // *Право. Журнал Высшей школы экономики*. – 2022. – № 1. – С. 4–27.
5. Харитонова Ю. С. Правовые средства обеспечения принципа прозрачности искусственного интеллекта // *Journal of Digital Technologies and Law*. – 2023. – Т. 1, № 2. – С. 337–358.
6. Alarcón A. A. The economics of language // *Catalan Sociolinguistics: State of the art and future challenges*. – 2022. – Т. 32. – С. 173.
7. Baugh J. Linguistic profiling across international geopolitical landscapes // *Dædalus*. – 2023. – Т. 152, № 3. – С. 167–177.
8. Bosco F. et al. Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities // *Reforming European data protection law*. – 2015. – С. 3–33.

9. Browning M., Arrigo B. Stop and risk: Policing, data, and the digital age of discrimination // *American Journal of Criminal Justice*. 2021. T. 46, № 2. С. 298–316.
10. Chasalow K., Levy K. Representativeness in statistics, politics, and machine learning // *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. – 2021. – С. 77–89.
11. Chilingaryan K., Meshkova I., Sheremetieva O. International legal protection of linguistic minorities // *International Journal of Psychosocial Rehabilitation*. – 2020. – T. 24, № 6. – С. 9750–9758.
12. Dash, N. S., & Arulmozi, S. History, features, and typology of language corpora. – Springer Singapore, 2018. – С. 1 – 291.
13. Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding // *arXiv preprint arXiv:1810.04805*. – 2018.
14. Doğruöz A. S., Sitaram S., Yong Z. X. Representativeness as a forgotten lesson for multilingual and code-switched data collection and preparation // *arXiv preprint arXiv:2310.20470*. – 2023. – С. 5751–5767.
15. Drożdżowicz A., Peled Y. The complexities of linguistic discrimination // *Philosophical Psychology*. – 2024. – С. 1–24.
16. Egbert J, Biber D, Gray B. Approaches to Representativeness in Previous Corpus Linguistic Research. In: *Designing and Evaluating Language Corpora: A Practical Framework for Corpus Representativeness*. Cambridge University Press. – 2022. – С. 28–51.
17. Engel C., Linhardt L., Schubert M. Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism // *Artificial Intelligence and Law*. – 2024. – С. 1–23.
18. Glauner P. Technical foundations of generative AI models. In: *Legal Tech – Zeitschrift für die digitale Anwendung des Nomos*. – С. 24–34.
19. Hacker P. A legal framework for AI training data—from first principles to the Artificial Intelligence Act // *Law, Innovation and Technology*. – 2021. – T. 13, № 2. – С. 257–301.
20. Hacker P. et al. Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It // *arXiv preprint arXiv:2407.10329*. – 2024. – С. 1–51
21. Ilya, I. Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data // *Legal Issues in the digital Age*. – 2024. – № 2. – С. 36–56.
22. Ilya I. The Voice and Speech Processing within Language Technology Applications: Perspective of the Russian Data Protection Law // *Legal Issues in the digital Age*. – 2020. – № 1. – С. 99–123.
23. Ilin I. Legal Regime of the Language Resources in the Context of the European Language Technology Development // *Language and Technology Conference*. – Cham: Springer International Publishing, 2019. – С. 367–376;
24. Ilin I., Kelli A. The use of human voice and speech in language technologies: the EU and Russian intellectual property law perspectives // *Juridica Int'l*. – 2019. – T. 28. – С. 17–27.

25. Ilin, I., & Kelli, A. Natural Language, Legal Hurdles: Navigating the Complexities in Natural Language Processing Development and Application. *The Journal of the University of Latvia. Law.* – C. 2–26 (forthcoming 2024).
26. Ilin I., Dedova M. Academic Entrepreneurship in the Field of Language Resource Creation and Dissemination. *The ESU 2019 Conference and Doctoral Programme, Naples (Italy), 8–14 September 2019.* Ed. Angelo Riviezzo, Maria Rosaria Napolitano, Antonella Garofano. *Electronic Conference Proceedings.* – C. 193–200.
27. Jiang X. et al. Large Language Models in Healthcare Current Development and Future Directions // *Generative AI Research.* – 2023. № 2. – C.12
28. Kelli A. et al. Constitutional values as a basis for the limitation of copyright within the context of digitalisation of the estonian language // *Constitutional Values in Contemporary Legal Space II.* – 2017. – C. 126 – 139.
29. Kome D. Legal protection of linguistic minority under discrimination: the case of anglophone Cameroon. – 2023. – C. 3–15
30. Li X. et al. Corpuslm: Towards a unified language model on corpus for knowledge-intensive tasks // *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval.* – 2024. – C. 26–37.
31. Möller J. T. Case Law of the UN Human Rights Committee relevant to Members of Minorities and Peoples in the Arctic Region // *The Yearbook of Polar Law Online.* – 2011. – T. 3, № 1. – C. 27–56.
32. Monteith B. et al. Unleashing the Economic Potential of Large Language Models: The Case of Chinese Language Efficiency // *Authorea Preprints.* – 2023. – C. 1–8.
33. Morin S. L. AI Discrimination in Hiring // *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology.* – IGI Global, 2024. – C. 64–74.
34. Orwat C. Algorithmic Discrimination From the Perspective of Human Dignity // *Social Inclusion.* – 2024. – T. 12. – C. 1–18.
35. Ozkul D. Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination // *Social Inclusion.* – 2024. – T. 12. C. 1–3.
36. Page C. Academic language development and linguistic discrimination: Perspectives from internationally educated students // *Comparative and International Education.* – 2023. – T. 52, № 2. – C. 39–53.
37. Rogers S. E. Bridging the 21st century digital divide // *TechTrends.* – 2016. – T. 60, № 3. – C. 197–199.
38. Sohail A., Zhang L. Using large language models to facilitate academic work in psychological sciences. – 2024. – C. 2–21.
39. Solovyev V. D., Akhtyamova S. Linguistic Big Data: Problem of Purity and Representativeness // *DAMDID/RCDL.* – 2019. – C. 193–204.
40. Vaswani A. Attention is all you need // *Advances in Neural Information Processing Systems.* – 2017.
41. Yenduri G. et al. Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions // *arXiv preprint arXiv:2305.10435.* – 2023.