Constitutional-Legal Aspect of Creating Large Language Models: the Problem of Digital Inequality and Linguistic Discrimination

Ilya G. Ilin

Saint Petersburg State University, Saint Petersburg, Russia

Keywords

artificial intelligence,
constitutional rights,
digital inequality,
digital technologies,
generative artificial
intelligence,
human rights,
large language models,
law,
linguistic discrimination,
natural language processing

Abstract

Objective: to study the impact of digital inequality on the implementation of constitutional human rights; to identify the risks of linguistic discrimination associated with the development and use of large language models.

Methods: formal-legal and comparative-legal methods, as well as the method of theoretical modeling. These approaches are complemented by general scientific methods of cognition, allowing for a comprehensive analysis of the legal, technological and social aspects of the issue.

Results: the research found that, in relation to large language models, digital inequality arises due to the uneven digitalization of languages and manifests itself in limited access to natural language processing technology. In turn, unequal access to this technology can negatively affect the implementation of constitutionally guaranteed rights and can be viewed from the viewpoint of equality and non-discrimination concepts. The author emphasizes that unequal access to natural language processing technologies can exacerbate existing social and economic inequalities and create new forms of discrimination.

Scientific novelty: hidden and indirect forms of discrimination are analyzed that manifest themselves in artificial intelligence systems, especially in generative models. While direct forms of discrimination can be detected in predictive algorithms, generative models create more subtle but no less significant cumulative effects. These effects contribute to the formation of social stereotypes and inequalities in areas such as professional activity, gender and ethnicity. The author also draws attention to the fact that with the increasing autonomy of artificial intelligence, traditional approaches

to discrimination detection are becoming less effective, which requires the development of new analysis and regulation methods.

Practical significance: the results provide a basis for identifying and assessing the legal risks associated with unequal access to digital products using natural language processing. This contributes to the improvement of legal regulation in the field of the development and use of artificial intelligence technologies. The article offers recommendations for lawmakers, regulators, and technology developers aimed at minimizing the risks of digital inequality and linguistic discrimination.

For citation

Ilin, I. G. (2025). Constitutional-Legal Aspect of Creating Large Language Models: the Problem of Digital Inequality and Linguistic Discrimination. *Journal of Digital Technologies and Law, 3*(1), 89–107. https://doi.org/10.21202/jdtl.2025.4

Contents

Introduction

Large Language Models (LLM) are generative artificial intelligence models used in natural language processing technology (NLP). They allow a computer to efficiently process text data, demonstrating the ability to "understand" text at a deep level, create coherent and contextually relevant responses to queries, translate from one language to another, and generate texts that meet certain stylistic and content requirements (Glauner, 2024). Examples of large language models include BERT[1], GPT-3[2], and related digital products such as Google Assistant and ChatGPT.

---

[1] Bidirectional Encoder Representations from Transformers (BERT) is a large language model developed by Alphabet Inc. (USA), based on the Transformer architecture. It is trained on a bidirectional context – it can analyze and "understand" text both from left to right and from right to left. For more information about the BERT model, see (Devlin J. et al., 2018).

[2] Generative Pre-trained Transformer (GPT) is a series of large language models developed by Openway (USA), based on the Transformer architecture. It is taught without a "teacher", does not require adaptation and can be used and adapted for a wide range of tasks. For more information about the GPT model, see (Yenduri G. et al., 2023). For more information about the Transformer architecture, see (Vaswani, 2017).

Large language models are trained on vast arrays of linguistic data, including structured linguistic corpora: databases containing a variety of texts (books, text transcriptions, translations, etc.) and audio files (audiobooks, broadcast recordings, podcasts, and other audio content). The structure and representativeness of such data, their volume and format determine the learning process and the accuracy of understanding the context (Ilin, 2024), while defects[3] or insufficient data can lead to incorrect model functioning and generally hinder the technology development (Hacker, 2021). Thus, the possibility of creating a high-quality language model directly depends on the volume, representativeness and other qualitative characteristics of the training data for a particular language.

At the same time, the levels of digitalization of languages – the volume of existing linguistic corpora and the data for their creation – differ significantly. For some languages or dialects, data may be extremely limited or non-existent. This hinders the development of accurate and effective language models, slowing down their digital development and limiting their integration into modern technologies. For example, if a data set is not comprehensive enough and does not cover all variants of a particular language, the model may process incoming requests incorrectly or inaccurately, and in some cases may not function at all. Differences in pronunciation, vocabulary, and grammar can lead to errors in recognizing and analyzing text or speech, and reduce the quality of results.

The inability to create a full-fledged language model for particular languages or dialects makes unavailable many digital products for speakers of these languages or significantly worsens the quality of their functioning compared to how the same technologies function for speakers of languages with a high level of digitalization. As a result, digital inequality arises, when access to modern technologies is unevenly distributed among different linguistic communities, which, in turn, increases the risk of discrimination.

The article aims to analyze the constitutional and legal aspects of creating large language models in the context of digital inequality and linguistic discrimination. To achieve this goal, we will investigate how digital inequality affects constitutional human rights, as well as analyze the risks of linguistic discrimination associated with the creation of large language models.

The article contains the main results of the corresponding research, as well as directions for further study. The research paper is divided into three thematic sections, supplemented by an introduction and conclusion. The first section analyzes the problem of digital inequality given the different levels of digitalization of languages – the volume and representativeness of linguistic data. The second section examines linguistic discrimination as a potential form of digital inequality, with an emphasis on unequal

---

[3]　In this context, a data defect includes both the data not meeting certain technical criteria and metrics, for example, the criteria of representativeness, volume, purity, etc. (quality defect) and a legal defect – the use of data in violation of the applicable legal regime. For example, a violation of the personal data regime during their processing as part of the language model. For more information about the impact of data quality on creating large language models, see (Ilin, 2024).

access to natural language processing (NLP). In the third section, the problem of linguistic discrimination is conceptualized in relation to other human rights and in the context of the digital technologies development.

1. Digitalization of languages as a source of digital inequality: a technical and legal analysis

Digital inequality is a form of social inequality characterized by unequal access to information technologies and by varying levels of skills in using them among individuals and social groups (Mushakov, 2022). This phenomenon encompasses a wide range of factors, including differences in technical equipment, access to Internet, digital literacy, and educational opportunities, which in turn leads to social and economic division (Rogers, 2016). The need to address gaps in access to digital technologies in order to achieve a more equal and inclusive society has been repeatedly highlighted both at the national[4] and international[5] levels.

In the context of creating large language models, the problem of digital inequality manifests itself in the limited ability of speakers of low-digitalization languages to use digital products in their language. This leads to unequal access of individuals or social groups to natural language processing (NLP) technologies. As a result, there may appear restrictions on access to information, education, and social services for native speakers of such languages. For example, the ability to understand text contextually and generate appropriate responses contributes to the active use of this technology in areas such as education and healthcare (Jiang et al., 2023; Sohail & Zhang, 2024). The lack of support for particular languages in these areas may negatively affect the exercising of the corresponding constitutionally guaranteed rights: the right to access education[6] and medical care[7]. This may limit the availability and quality of these services. In this regard, it seems logical to consider the problem of digital inequality from the viewpoint of constitutional and legal relations, i.e. the concepts of "equality" and prohibition of discrimination[8].

---

[4] Decree of the Government of the Russian Federation No. 313 dated 15.04.2014 (2014). Here and further, all references to documents, regulations and judicial practice are given by SPS ConsultantPlyus refence system. https://clck.ru/3GP8do

[5] Geneva Declaration of Principles (Building the Information Society: A Global Challenge in the New Millennium) (UN) of December 12, 2003 https://clck.ru/3GP8fD ; Tunisian Programme for the Information Society (UN) dated November 15, 2005. https://clck.ru/3GP8ge

[6] The Constitution of the Russian Federation, adopted by popular vote on 12.12.1993 with amendments approved during the nationwide vote on 01.07.2020 (hereinafter referred to as the Constitution of the Russian Federation). Art. 43. https://clck.ru/3GP8hh

[7] Constitution of the Russian Federation. Art. 41. https://clck.ru/3GP8jK

[8] In both cases, the issue of equality of rights is considered, but the right to non-discrimination has a narrower content and in this sense stems from the general right to equality. For more information, see (Talapina, 2022).

We can also agree with some researchers that constitutional norms ensuring equality before the law and access to services should take into account and eliminate inequality in access to digital resources, as this directly affects the ability of citizens to exercise their rights and freedoms in the digital era (Mushakov, 2022).

To create an effective language model, a set of training data is needed that must meet criteria such as volume, representativeness[9], and other qualitative characteristics. These parameters directly depend on the level of digitalization of a particular language, since the higher the degree of digitalization, the more diverse and high-quality data can be used to train the model.

Digitalization of a language in a broad sense is the transformation of data into appropriate electronic linguistic corpora. For this purpose, text data (for example, files, transcriptions, abstracts), speech data (for example, audio recordings, phonetic and intonation abstracts) and multimodal data (i.e. data combining several types, for example, video and text, images and text, etc.) are used (Dash & Arulmozi, 2018). It should be noted that this process not only contributes to technological development and digital transformation of society, but also plays an important role in preserving national and cultural identity (Kelli et al., 2016). For example, digitalization of minority languages can significantly contribute to the preservation of the cultural heritage of small nations.

Despite the importance of digitalization for technological progress and the high social significance of this process, the level of digitalization of languages and their dialects remains uneven. There are economic, technical, and legal factors that limit or hinder the digitalization of languages.

The economic factors are related to the fact that languages have different economic potential (Alarcón, 2022; Monteith & Sung, 2023). Hence, digitalization requires significant resources, including time, finance, etc. In this regard, the development of linguistic corpora for some languages may be economically unfeasible. The technical factors are directly related to creating linguistic corpora. Such factors may include errors in data collection, flaws in the corpora design and limitations of existing datasets, errors in metadata, etc. (Solovyev & Akhtyamova, 2019; Doğruöz et al., 2023; Li et al., 2024). The legal factors are related to the presence of regulatory restrictions on access to training data and the need to comply with the relevant legal regime when using them for training.

In previous works, the author discussed in detail the issues of regulating access to training data (Ilin, 2024), as well as compliance with their legal regimes, such as the personal data regime (Ilin, 2020) and the intellectual property regime

---

[9] Given the multifaceted meaning of the term "representativeness" (for more details see (Chasalow & Levy, 2021)), it is important to note that in the context of this article, the volume of linguistic data means their quantity, while representativeness means their diversity, i.e. the degree to which various styles, dialects, time periods and contexts are covered.

(Ilin, 2022; Ilin & Kelli, 2019, 2024). The central topic of those studies was the conflict between equally protected human rights when using training data, such as the right to non-discrimination[10] and the right to privacy, to protection of personal and family secrets[11]. Overcoming this problem is necessary both at the conceptual level (removing regulatory barriers to data access, taking into account the balance of private and public interests) and in practical terms (creating conditions for the dissemination and exchange of linguistic data, for example, by developing an institution of reusing the data accumulated in government information systems (hereinafter referred to as GIS) or involving higher educational institutions to create linguistic corpora and digitalize the language).

According to the analytical report of the Accounts Chamber of the Russian Federation[12], by 2020, more than 800 federal state information systems were operating in Russia, providing data exchange between government agencies in various areas of public life. These systems cover a wide range of information, including statistics, as well as information on healthcare, education, and other key sectors. In this context, the use of GIS data to create linguistic corpora seems to be particularly promising. Despite the varying levels of development of these systems, one may expect that the data collected in them will have the necessary qualitative characteristics, and their diversity can provide the necessary representativeness and volume (Ilin, 2024). However, given the risks associated with legal restrictions on the use of data, their reuse should be carried out in accordance with uniform principles and regulations. These should include legislative standards and control mechanisms that take into account the specifics of each data type and its compliance with the purposes of its initial collection.

Another possible solution to the problem of access and lack of linguistic data is to involve higher education institutions to creating and subsequently disseminating linguistic corpora. The participation of universities in language digitalization can also be justified by the social significance of this activity. As an example of successful cooperation between commercial organizations and higher education institutions in the field of natural language processing, we can mention the joint academic program of the Center for Speech Technologies Group of Companies with the ITMO National Research University (Ilin & Dedova, 2019).

However, although this solves the problem of creating linguistic corpora, the issue of their further distribution remains open. For example, for various reasons, a university may not be interested in further dissemination of the linguistic corpus or may not have

---

[10] Constitution of the Russian Federation. Art. 19. https://clck.ru/3GPBg6

[11] Constitution of the Russian Federation. Art. 23. https://clck.ru/3GPBhb

[12] Center for Advanced Governance (2020). Assessment of the openness of government information systems in Russia: analytical report. https://clck.ru/3GPBjT

the necessary resources for this and, accordingly, may not distribute it. If a university operates as an entrepreneurial university and commercializes its results, for example, via a spin-off company, one may also question the possibility to employ the doctrine of the free use of works[13] when processing linguistic data. All these questions require further careful analysis, both from a legal and other points of view.

2. Linguistic discrimination as a form of digital inequality

Since, in relation to large language models development, digital inequality leads to unequal access of individuals or social groups to natural language processing (NLP) technologies (the inability to fully use this technology in their language), the problem of digital inequality should primarily be considered in the context of linguistic discrimination.

The problem of discrimination by artificial intelligence systems, although not new, remains relevant today. The development and active implementation of artificial intelligence systems in various spheres of life opens up new areas for discussion of this problem. Examples include discrimination by artificial intelligence systems in the field of labor relations (Morin, 2024), the impact of profiling[14] on human dignity (Orwat, 2024), the potential impact of artificial intelligence on discrimination based on ethnicity, religion and gender (Ozkul, 2024), etc.

In addition, with the increasing autonomy of artificial intelligence systems and the development of generative models, discrimination begins to take on an implicit character, which allows classifying its manifestations into direct and indirect ones. For example, unlike the obvious cases of discrimination observed in predictive crime analytics systems such as those based on PredPol[15] and COMPAS[16] algorithms, discrimination in generative artificial intelligence systems may be less apparent. For example, these systems may preferentially create images of white men in response

---

[13] Civil Code of the Russian Federation (part 4) of 18.12.2006 No. 230-FZ. Art. 1274. https://clck.ru/3GPBnL

[14] Profiling is a technique of intellectual data analysis that can be automated or semi-automated and aims to create classes or categories of characteristics from large datasets. In this process, data is collected, analyzed using various algorithms such as machine learning, and used to create profiles describing typical characteristics or behavioral patterns of groups or individuals. For more information, see (Bosco et al., 2015).

[15] PredPol (Predictive Policing) is a predictive analytics system used by police and designed to predict crimes. PredPol's main goal is to use historical data on crime to create maps of "hot spots" – areas where crimes are most likely to occur. For more information, see (Browning & Arrigo, 2021).

[16] COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a predictive analytics system designed to assess the risk of recidivism among convicts. Its main goal is to analyze data on offenses, behavior, and the social record of suspects in order to predict the likelihood of their reoffending. The system is used in judicial practice to help making decisions on sentencing and release conditions. For more information, see (Engel et al., 2024).

to repeated requests for examples of people employed in important professions, potentially leading to cumulative discriminatory effects (Hacker et al., 2024). In such cases, discrimination becomes difficult to detect, as it may not be explicit or obvious, but nevertheless has a significant impact on the representation and perception of various groups in society.

The National Strategy for the Development of Artificial Intelligence up to 2030[17] (hereinafter referred to as the Strategy) emphasizes that the protection of human rights and freedoms is one of the main principles of the development and use of artificial intelligence technology[18], while "non-discrimination" is highlighted as one of the main principles of the development of legal regulation of public relations in the sphere of development and use of artificial intelligence technology[19].

Article 2 of the Universal Declaration of Human Rights (1948)[20] prohibits discrimination, including on the basis of language. A similar provision is contained in Article 1 (3) of the UN Charter[21], and is also reflected in paragraph 2 of Article 19 of the Constitution of the Russian Federation, according to which the state guarantees equality of human and civil rights and freedoms, regardless of language.

In the field of language discrimination, several key aspects are identified, related to its recognition, legal protection and public perception. One of the main problems is the lack of recognition of linguistic discrimination at the international level. For example, discrimination based on voice often goes unnoticed (Baugh, 2023), which can be critical when interacting with speech and voice recognition technology and related digital products: interactive response systems and voice assistants.

The UN Human Rights Committee[22] has repeatedly addressed the issue of linguistic discrimination, but its judicial practice is underdeveloped and does not provide reliable protection for linguistic minorities (Möller, 2011). The regulatory framework at various

---

[17] The National Strategy for the Development of Artificial Intelligence up to 2030 was approved by Decree of the Russian President "ated 10.10.2019 No. 490 «On the development of artificial intelligence in the Russian Federation" (hereinafter referred to as the National Strategy for the Development of Artificial Intelligence up to 2030).

[18] National Strategy for the Development of Artificial Intelligence up to 2030. 19 (a). https://clck.ru/3Ghyfz

[19] National Strategy for the Development of Artificial Intelligence up to 2030. cl. 51 (10) (d). https://clck.ru/3Ghyfz

[20] The Universal Declaration of Human Rights (adopted by the UN General Assembly on 10.12.1948). https://clck.ru/3GPBqd

[21] Charter of the United Nations Organization (adopted in San Francisco on 26.06.1945). https://clck.ru/3GPBsN

[22] The UN Human Rights Committee was established on the basis of the International Covenant on Civil and Political Rights, which was adopted by the UN General Assembly in 1966 and entered into force in 1976. The Committee is the body that monitors the fulfillment of the obligations assumed under this Covenant by participating states. The Committee considers reports from the states on how they respect the rights enshrined in the Covenant, as well as individual complaints about violations of rights (if the state has recognized the Committee's jurisdiction in this matter). More about the Committee: https://clck.ru/3GPBvJ

levels also often does not take into account all the nuances of linguistic discrimination. Legislation at the international, regional and national levels, as a rule, does not provide sufficient protection for the rights of linguistic minorities, which leads to gaps in the legal protection of crime victims (Chilingaryan et al., 2020).

Discrimination based on language can be defined as any unjustified distinction or restriction that weakens or excludes the possibility of exercising rights enshrined in international or national regulations based on language affiliation. At the same time, it should be added that states also have positive obligations to protect and promote linguistic rights as part of their obligation to respect human rights[23]. Therefore, in the context of creating large language models, it seems necessary to expand the definition of linguistic discrimination to include actions aimed at hindering the preservation or development of minority languages. The essence of the first part of the definition is that linguistic discrimination occurs when a person experiences worse treatment than others in a similar situation due to insufficient or complete lack of proficiency in the official language established in a given state or region. The second part refers to a deeper aspect of this problem – the states' fulfillment of legal obligations, established by international conventions and national legislation, to protect and promote minority languages. This said, it should be noted that the expansion of the linguistic discrimination concept is more likely to reflect the perspective sought by judicial practice and scientific discussion, rather than the current perception of the problem by law enforcers and lawyers.

3. Qualification problem and criteria for assessing linguistic discrimination

Ambiguity in the definition of linguistic discrimination makes law enforcement difficult and raises questions about the criteria used in assessing these situations. As noted earlier, linguistic discrimination occurs when people are treated differently because of their language proficiency or accent, which often leads to limited access to opportunities and rights (Mironova, 2019). However, linguistic discrimination is a multifaceted problem that differs from other forms of discrimination, such as racial or religious, and depends on various factors. An analysis of existing practice allows us to identify a number of key factors for determining linguistic discrimination. The first factor is the number of native speakers: the level of discrimination is often determined by the prevalence of language in society. For example, in Cameroon, the English-speaking minority faces systemic discrimination due to its small size compared to the French-speaking majority (Donard, 2023).

---

[23] For example, the obligations arising from Federal Law No. 273-FZ of December 29, 2012 "On education in the Russian Federation", Federal Law No. 74-FZ of 17.06.1996 "On national cultural autonomy".

Another important factor is the ability of the state to support multilingualism. The more actively the state creates conditions for learning and using multiple languages, the lower the likelihood of linguistic discrimination. For example, research shows that support for multilingualism in educational institutions helps to reduce discrimination based on language (Page, 2023).

The use of minority languages in public life is also of great importance. When these languages do not receive institutional support, their speakers are often marginalized, which reinforces existing social inequalities.

In addition, it should be borne in mind that linguistic discrimination may overlap with other forms of discrimination, such as racial, religious, or ethnic. In such cases, people are subjected to complex forms of discrimination, which significantly exacerbates the problem (Drożdżowicz & Peled, 2024). In order to illustrate the complexity of the problem, let us briefly consider some of these intersections.

Failure to provide equal access to services in the mother tongue may violate the right to equality, creating barriers that hinder full participation in social life[24]. These barriers, for example, can affect the right to education[25] by limiting access to educational resources and materials in the mother tongue, which can reduce the quality of education and limit educational opportunities.

In addition, linguistic discrimination affects the right to freedom of expression[26]. People should be able to express their opinions freely in the language they prefer, and restrictions on this may be seen as a violation of this fundamental right. Linguistic discrimination also affects cultural rights, as language is a key element of cultural identity and expression. Restricting the use of a minority language in cultural and social contexts can undermine the cultural rights of these communities and their ability to preserve and develop their cultural identity.

Access to justice can also be hampered by language barriers, as the need to understand and participate in court proceedings in one's native language is critical to ensuring fair justice[27]. Language barriers may prevent the correct understanding of charges, court procedure, or legal decisions, which can lead to unfair outcomes.

---

[24] D.H. and Others v. Czech Republic: Judgment of the Grand Chamber of the European Court of Human Rights of November 13, 2007 (complaint No. 57325/00).

[25] Communication No. 760/1997. J.G.A. Diergaardt (late Captain of the Rehoboth Baster Community) et al. v. Namibia, Views of 25 July 2000, CCPR/C/69/D/760/1997.

[26] Communication No. 221/1987. Yves Cadoret and Hervé Le Bihan v. France, Views of 11 April 1991, CCPR/C/41/D/221/1987; Communication No. 219/1986. Dominique Guesdon v. France, Views of 25 July 1990, CCPR/C/39/D/219/1986.

[27] For example, the court's refusal to provide the accused with the text of the indictment translated into the Karachay language led to the cancellation of the verdict due to violations of the norms of criminal and criminal-procedural law by the preliminary investigation authorities. For more information, see "Review of the cassation practice of the Judicial Board for Criminal Cases of the Supreme Court of the Russian Federation of 2003" (2004). Bulletin of the Supreme Court of the Russian Federation, 9.

Thus, although it is possible to identify factors for assessing linguistic discrimination, the legal qualification of such cases in the context of digital technologies causes certain difficulties. For example, it is necessary to find out whether errors in the language model can be considered a manifestation of discrimination. Such errors are often difficult to detect, as discrimination may be hidden, which makes it less obvious for analysis. Discrimination in models can be the result of algorithmic or human bias. Algorithmic bias occurs due to limitations or distortions in the data on which the model is trained, whereas human bias can manifest itself in the developing and configuring algorithms (Kharitonova et al., 2021). Both forms of bias can not only affect the accuracy and fairness of decisions, but also maintain or exacerbate existing social inequalities, ultimately leading to discrimination. The distinction between errors and discrimination requires in-depth analysis, as errors may be accidental or may result from systemic biases. It is important to understand how bias, both algorithmic and human, affects the decision-making process and how it is integrated into algorithms and models. This understanding is necessary to develop more equitable and inclusive digital systems.

Conclusions

This article aims to analyze the constitutional-legal aspects of creating large language models in the context of digital inequality and linguistic discrimination. The study found that digital inequality in the context of large language models is due to the uneven digitalization of languages and manifests itself in limited access to natural language processing technologies. Such unequal access can negatively affect the implementation of constitutionally guaranteed rights and requires consideration through the prism of such concepts as "equality" and prohibition of discrimination. In turn, the identification and legal qualification of linguistic discrimination when creating large language models is a difficult task, since biases in models can be hidden and have a cumulative discriminatory effect. Discrimination may be caused by both algorithmic and human bias. Algorithmic bias occurs due to limitations or distortions in the data on which the model is trained, while human bias can manifest itself in developing and configuring algorithms. Distinguishing between these categories and assessing their impact on decision-making are becoming important areas for future research aimed at developing mechanisms to ensure equal access to digital technologies and the protection of language rights.

References

Alarcón, A. A. (2022). The economics of language. In Miquel Àngel Pradilla Cardona (Ed.), *Catalan Sociolinguistics: State of the art and future challenges* (pp. 173–182). https://doi.org/10.1075/ivitra.32.12ala

Baugh, J. (2023). Linguistic profiling across international geopolitical landscapes. *Daedalus*, *152*(3), 167–177. https://doi.org/10.1162/daed_a_02024

Bosco, F., Creemers, N., Ferraris, V., Guagnin, D., & Koops, B. J. (2015). Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. In S. Gutwirth, R. Leenes, P. de Hert (Eds.), *Reforming European data protection law* (Vol. 20, pp. 3–33). https://doi.org/10.1007/978-94-017-9385-8_1

Browning, M., & Arrigo, B. (2021). Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice*, *46*(2), 298–316. https://doi.org/10.1007/s12103-020-09557-x

Chasalow, K., & Levy, K. (2021). Representativeness in statistics, politics, and machine learning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 77–89). https://doi.org/10.1145/3442188.3445872

Chilingaryan, K., Meshkova, I., & Sheremetieva, O. (2020). International legal protection of linguistic minorities. *International Journal of Psychosocial Rehabilitation*, *24*(6), 9750–9758. https://doi.org/10.37200/IJPR/V24I6/PR26097

Dash, N. S., & Arulmozi, S. (2018). *History, features, and typology of language corpora*. Springer Singapore. https://doi.org/10.1007/978-981-10-7458-5

Devlin, J., Chang, Ming-Wei, Lee, Kenton, & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Doğruöz, A. S., Sitaram, S., & Yong, Z. X. (2023). *Representativeness as a forgotten lesson for multilingual and code-switched data collection and preparation*. arXiv preprint arXiv:2310.20470 (pp. 5751–5767).

Donard, K. (2023). Legal protection of linguistic minority under discrimination: the case of anglophone Cameroon. *International Journal of Business and Technology*, *11*(2), Article 1.

Drożdżowicz, A., & Peled, Y. (2024). The complexities of linguistic discrimination. *Philosophical Psychology*, *37*(6), 1459–1482. https://doi.org/10.1080/09515089.2024.2307993

Engel, C., Linhardt, L., & Schubert, M. (2024). Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism. In *Artificial Intelligence and Law*. https://doi.org/10.1007/s10506-024-09389-8

Glauner, P. (2024). Technical foundations of generative AI models. In *Legal Tech – Zeitschrift für die digitale Anwendung*, *1*, 24–34.

Hacker, P. A (2021). Legal framework for AI training data—from first principles to the Artificial Intelligence Act. *Law, Innovation and Technology*, *13*(2), 257–301. https://doi.org/10.1080/17579961.2021.1977219

Hacker, P., Mittelstadt, B., Zuiderveen Borgesius, F., Wachteret, S. (2024). *Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It*. arXiv preprint arXiv:2407.10329. https://doi.org/10.2139/ssrn.4877398

Ilin, I. (2020). The Voice and Speech Processing within Language Technology Applications: Perspective of the Russian Data Protection Law. *Legal Issues in the Digital Age*, *1*, 99–123. https://doi.org/10.17323/2713-2749.2020.1.99.123

Ilin, I. (2022). Legal Regime of the Language Resources in the Context of the European Language Technology Development. In Z. Vetulani, P. Paroubek, M. Kubis (Eds.), *Human Language Technology. Challenges for Computer Science and Linguistics. LTC 2019. Lecture Notes in Computer Science* (vol. 13212, pp. 367–376). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-05328-3_24

Ilin, I. (2024). Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data. *Legal Issues in the Digital Age*, *2*, 36–56. https://doi.org/10.17323/2713-2749.2024.2.36.56

Ilin, I., & Dedova, M. (2019). Academic Entrepreneurship in the Field of Language Resource Creation and Dissemination. In A. Riviezzo, M. Rosaria Napolitano, & A. Garofano (Eds.), *The ESU 2019 Conference and Doctoral Programme, Naples (Italy), 8–14 September 2019. Electronic Conference Proceedings* (pp. 193–200).

Ilin, I., & Kelli, A. (2019). The use of human voice and speech in language technologies: the EU and Russian intellectual property law perspectives. *Juridical International*, *28*, 17–27. https://doi.org/10.12697/ji.2019.28.03

Ilin, I., & Kelli, A. (2024). Natural Language, Legal Hurdles: Navigating the Complexities in Natural Language Processing Development and Application. *Journal of the University of Latvia. Law*, *17*, 44–67. https://doi.org/10.22364/jull.17.03

Jiang, X., Yan, L., Vavekanand, R., & Hu, M. (2023). Large Language Models in Healthcare Current Development and Future Directions. *Generative AI Research*, *2*, 12. https://doi.org/10.20944/preprints202407.0923.v1

Kelli, A., Vider, K., Pisuke, H., & Siil, T. (2016). Constitutional values as a basis for the limitation of copyright within the context of digitalisation of the Estonian language. In *Constitutional Values in Contemporary Legal Space* (Vol. II, pp. 126–139).

Kharitonova, Yu. S., Savina, V. S., & Pagnini, F. (2021). Artificial Intelligence's Algorithmic Bias: Ethical and Legal Issues. Perm University Herald. *Juridical Sciences*, *53*, 488–515. (In Russ.). https://doi.org/10.17072/1995-4190-2021-53-488-515

Li, X., Dou, Zh., Zhou, Yu., & Liu, F. (2024). CorpusLM: Towards a unified language model on corpus for knowledge-intensive tasks. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 26–37). https://doi.org/10.1145/3626772.3657778

Mironova, M. V. (2019). Formation of the term "Linguistic discrimination" in modern sociolinguistics. In *New Language. New World. New Thinking: collection of works of the 2nd Annual international scientific-practical conference* (pp. 555–558). Moscow: Diplomatic Academy of Ministry of Foreign Affairs of the Russian Federation. (In Russ.).

Möller, J. T. (2011). 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*, *3*(1), 27–56. https://doi.org/10.1163/22116427-91000054

Monteith, B., & Sung, M. (2023). Unleashing the Economic Potential of Large Language Models: The Case of Chinese Language Efficiency. *TechRxiv*. *June 07*. https://doi.org/10.36227/techrxiv.23291831.v1

Morin, S. L. (2024). AI Discrimination in Hiring. In D. Norman (Ed.), *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology* (pp. 64–74). IGI Global. https://doi.org/10.4018/979-8-3693-1906-2.ch004

Mushakov, V. (2022). Constitutional human rights in the context of bridging the digital divide. *Vestnik of the St. Petersburg University of the Ministry of Internal Affairs of Russia*, *2022*(1). (In Russ.). https://doi.org/10.35750/2071-8284-2022-1-69-73

Orwat, C. (2024). Algorithmic Discrimination From the Perspective of Human Dignity. *Social Inclusion*, *12*, 1–18. https://doi.org/10.17645/si.7160

Ozkul, D. (2024). Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination. *Social Inclusion*, *12*, 1–3. https://doi.org/10.17645/si.8942

Page, C. (2023). Academic language development and linguistic discrimination: Perspectives from internationally educated students. *Comparative and International Education*, *52*(2), 39–53. https://doi.org/10.5206/cie-eci.v52i2.15000

Rogers, S. E. (2016). Bridging the 21st century digital divide. *TechTrends*, *60*(3), 197–199. https://doi.org/10.1007/s11528-016-0057-0

Sohail, A., & Zhang, L. (2024). *Integrating large language models into the psychological sciences*. https://doi.org/10.1007/s12144-025-07438-2

Solovyev, V. D., & Akhtyamova, S. (2019). Linguistic Big Data: Problem of Purity and Representativeness. In *21st International Conference on Data analytics and management in data intensive domains, DAMDID/RCDL 2019* (pp. 193–204).

Talapina, E. (2022). Artificial Intelligence Processing and Risks of Discrimination. *Law Journal of the Higher School of Economics*, *1*, 4–27. (In Russ.). https://doi.org/10.17323/2072-8166.2022.1.4.27

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. https://doi.org/10.48550/arXiv.1706.03762

Yenduri, G., Ramalingam, M., Chemmalar Selvi, G., Supriya, Y., Srivastava, G., Maddikunta, P. K. R. et al. (2023). Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. In *IEEE Access* (Vol. 12, pp. 54608–54649). https://doi.org/10.1109/access.2024.3389497

Author information

Ilya G. Ilin – Master of Law (information technologies), postgraduate student, Faculty of Law, Saint Petersburg State University
Address: 22nd line of Vasilievsky Island, 7199106 Saint Petersburg, Russian Federation
E-mail: i.g.ilin@spbu.ru
ORCID ID: https://orcid.org/0000-0003-1076-2765
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57765898000
WoS Researcher ID: https://www.webofscience.com/wos/author/record/FDF-0979-2022
Google Scholar ID: https://scholar.google.com/citations?user=YruuMK0AAAAJ
RSCI Author ID: https://elibrary.ru/author_profile.asp?authorid=1253542

Conflict of interest

The author declares no conflict of interest.

Financial disclosure

The research had no sponsorship.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – November 15, 2024
Date of approval – November 25, 2024
Date of acceptance – March 25, 2025
Date of online placement – March 30, 2025

Конституционно-правовой аспект создания больших языковых моделей: проблема цифрового неравенства и языковой дискриминации

Илья Геннадьевич Ильин

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Ключевые слова

большие языковые модели,
генеративный
искусственный интеллект,
искусственный интеллект,
конституционные права,
обработка естественного
языка,
права человека,
право,
цифровое неравенство,
цифровые технологии,
языковая дискриминация

Аннотация

Цель: исследование влияния цифрового неравенства на реализацию конституционных прав человека, а также выявление рисков языковой дискриминации, связанных с разработкой и использованием больших языковых моделей.

Методы: формально-юридический и сравнительно-правовой методы, а также метод теоретического моделирования. Эти подходы дополняются общенаучными методами познания, что позволяет провести комплексный анализ правовых, технологических и социальных аспектов проблемы.

Результаты: 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. В свою очередь, неравный доступ к указанной технологии может негативно влиять на реализацию конституционно гарантированных прав и может быть рассмотрен с точки зрения концепций «равенства» и запрета на дискриминацию. Автор подчеркивает, что неравный доступ к технологиям обработки естественного языка может усугублять существующие социальные и экономические неравенства, создавая новые формы дискриминации.

Научная новизна: заключается в анализе скрытых и косвенных форм дискриминации, которые проявляются в системах искусственного интеллекта, особенно в генеративных моделях. В отличие от прямых форм дискриминации, которые могут быть выявлены в предсказательных алгоритмах, генеративные модели создают более тонкие, но не менее значимые кумулятивные эффекты. Эти эффекты способствуют формированию социальных стереотипов и неравенства в таких областях, как профессиональная деятельность, гендерная и этническая принадлежность. Автор также обращает внимание на то, что с увеличением

автономности искусственного интеллекта традиционные подходы к выявлению дискриминации становятся менее эффективными, что требует разработки новых методов анализа и регулирования.

Практическая значимость: состоит в том в том, что его результаты предоставляют основу для выявления и оценки правовых рисков, связанных с неравным доступом к цифровым продуктам, использующим технологии обработки естественного языка. Это способствует совершенствованию правового регулирования в сфере разработки и использования технологий искусственного интеллекта. Статья предлагает рекомендации для законодателей, регулирующих органов и разработчиков технологий, направленные на минимизацию рисков цифрового неравенства и языковой дискриминации.

Для цитирования

Ильин, И. Г. (2025). Конституционно-правовой аспект создания больших языковых моделей: проблема цифрового неравенства и языковой дискриминации. *Journal of Digital Technologies and Law*, *3*(1), 89–107. https://doi.org/10.21202/jdtl.2025.4

Список литературы

Миронова, М. В. (2019). Становление термина «языковая дискриминация» в современной социолингвистике. В сб. *New Language. New World. New Thinking: сборник материалов II Ежегодной международной научно-практической конференции* (с. 555–558). Москва: Дипломатическая академия Министерства иностранных дел Российской Федерации. https://elibrary.ru/bjegvs

Мушаков, В. Е. (2022). Конституционные права человека в контексте проблемы преодоления цифрового разрыва. *Вестник Санкт-Петербургского университета МВД России*, *1*(93), 69–73. EDN: https://elibrary.ru/elrbud. DOI: https://doi.org/10.35750/2071-8284-2022-1-69-73

Талапина, Э. В. (2022). Обработка данных при помощи искусственного интеллекта и риски дискриминации. Право. *Журнал Высшей школы экономики*, *1*, 4–27. EDN: https://elibrary.ru/pwepsj. DOI: https://doi.org/10.17323/2072-8166.2022.1.4.27

Харитонова, Ю. С., Савина, В. С., Паньини, Ф. (2021). Предвзятость алгоритмов искусственного интеллекта: вопросы этики и права. *Вестник Пермского университета. Юридические науки*, *53*, 488–515. EDN: https://elibrary.ru/eukcpy. DOI: https://doi.org/10.17072/1995-4190-2021-53-488-515

Alarcón, A. A. (2022). The economics of language. In Miquel Àngel Pradilla Cardona (Ed.), *Catalan Sociolinguistics: State of the art and future challenges* (pp. 173–182). https://doi.org/10.1075/ivitra.32.12ala

Baugh, J. (2023). Linguistic profiling across international geopolitical landscapes. *Daedalus*, *152*(3), 167–177. https://doi.org/10.1162/daed_a_02024

Bosco, F., Creemers, N., Ferraris, V., Guagnin, D., & Koops, B. J. (2015). Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. In S. Gutwirth, R. Leenes, P. de Hert (Eds.), *Reforming European data protection law* (Vol. 20, pp. 3–33). https://doi.org/10.1007/978-94-017-9385-8_1

Browning, M., & Arrigo, B. (2021). Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice*, *46*(2), 298–316. https://doi.org/10.1007/s12103-020-09557-x

Chasalow, K., & Levy, K. (2021). Representativeness in statistics, politics, and machine learning. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 77–89). https://doi.org/10.1145/3442188.3445872

Chilingaryan, K., Meshkova, I., & Sheremetieva, O. (2020). International legal protection of linguistic minorities. *International Journal of Psychosocial Rehabilitation*, *24*(6), 9750–9758. EDN: https://elibrary.ru/dgcwtx. DOI: https://doi.org/10.37200/IJPR/V24I6/PR26097

Dash, N. S., & Arulmozi, S. (2018). *History, features, and typology of language corpora*. Springer Singapore. https://doi.org/10.1007/978-981-10-7458-5

Devlin, J., Chang, Ming-Wei, Lee, Kenton, & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv preprint arXiv:1810.04805.

Doğruöz, A. S., Sitaram, S., & Yong, Z. X. (2023). *Representativeness as a forgotten lesson for multilingual and code-switched data collection and preparation*. arXiv preprint arXiv:2310.20470 (pp. 5751–5767).

Donard, K. (2023). Legal protection of linguistic minority under discrimination: the case of anglophone Cameroon. *International Journal of Business and Technology*, *11*(2), Article 1.

Drożdżowicz, A., & Peled, Y. (2024). The complexities of linguistic discrimination. *Philosophical Psychology*, *37*(6), 1459–1482. https://doi.org/10.1080/09515089.2024.2307993

Engel, C., Linhardt, L., & Schubert, M. (2024). Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism. In *Artificial Intelligence and Law*. https://doi.org/10.1007/s10506-024-09389-8

Glauner, P. (2024). Technical foundations of generative AI models. In *Legal Tech-Zeitschrift für die digitale Anwendung*, *1*, 24–34.

Hacker, P. A (2021). Legal framework for AI training data—from first principles to the Artificial Intelligence Act. *Law, Innovation and Technology*, *13*(2), 257–301. https://doi.org/10.1080/17579961.2021.1977219

Hacker, P., Mittelstadt, B., Zuiderveen Borgesius, F., Wachteret, S. (2024). *Generative Discrimination: What Happens When Generative AI Exhibits Bias, and What Can Be Done About It*. arXiv preprint arXiv:2407.10329. https://doi.org/10.2139/ssrn.4877398

Ilin, I. (2022). Legal Regime of the Language Resources in the Context of the European Language Technology Development. In Z. Vetulani, P. Paroubek, M. Kubis (Eds.), *Human Language Technology. Challenges for Computer Science and Linguistics. LTC 2019. Lecture Notes in Computer Science* (vol. 13212, pp. 367–376). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-05328-3_24

Ilin, I., & Dedova, M. (2019). Academic Entrepreneurship in the Field of Language Resource Creation and Dissemination. In A. Riviezzo, M. Rosaria Napolitano, & A. Garofano (Eds.), *The ESU 2019 Conference and Doctoral Programme, Naples (Italy), 8–14 September 2019. Electronic Conference Proceedings* (pp. 193–200).

Ilin, I., & Kelli, A. (2024). Natural Language, Legal Hurdles: Navigating the Complexities in Natural Language Processing Development and Application. *Journal of the University of Latvia. Law*, *17*, 44–67. https://doi.org/10.22364/jull.17.03

Ilin, I., & Kelli, A. (2019). The use of human voice and speech in language technologies: the EU and Russian intellectual property law perspectives. *Juridical International*, *28*, 17–27. https://doi.org/10.12697/ji.2019.28.03

Ilin, I. (2020). The Voice and Speech Processing within Language Technology Applications: Perspective of the Russian Data Protection Law. *Legal Issues in the Digital Age*, *1*, 99–123. EDN: https://elibrary.ru/axbzzq. DOI: https://doi.org/10.17323/2713-2749.2020.1.99.123

Ilin, I. (2024). Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data. *Legal Issues in the Digital Age*, *2*, 36–56. EDN: https://elibrary.ru/azkzba. DOI: https://doi.org/10.17323/2713-2749.2024.2.36.56

Jiang, X., Yan, L., Vavekanand, R., & Hu, M. (2023). Large Language Models in Healthcare Current Development and Future Directions. *Generative AI Research*, *2*, 12. https://doi.org/10.20944/preprints202407.0923.v1

Kelli, A., Vider, K., Pisuke, H., & Siil, T. (2016). Constitutional values as a basis for the limitation of copyright within the context of digitalisation of the Estonian language. In *Constitutional Values in Contemporary Legal Space* (Vol. II, pp. 126–139).

Li, X., Dou, Zh., Zhou, Yu., & Liu, F. (2024). CorpusLM: Towards a unified language model on corpus for knowledge-intensive tasks. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 26–37). https://doi.org/10.1145/3626772.3657778

Möller, J. T. (2011). 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*, *3*(1), 27–56. https://doi.org/10.1163/22116427-91000054

Monteith, B., & Sung, M. (2023). Unleashing the Economic Potential of Large Language Models: The Case of Chinese Language Efficiency. *TechRxiv*. *June 07*. https://doi.org/10.36227/techrxiv.23291831.v1

Morin, S. L. (2024). AI Discrimination in Hiring. In D. Norman (Ed.), *Innovations, Securities, and Case Studies Across Healthcare, Business, and Technology* (pp. 64–74). IGI Global. https://doi.org/10.4018/979-8-3693-1906-2.ch004

Orwat, C. (2024). Algorithmic Discrimination From the Perspective of Human Dignity. *Social Inclusion*, *12*, 1–18. https://doi.org/10.17645/si.7160

Ozkul, D. (2024). Artificial Intelligence and Ethnic, Religious, and Gender-Based Discrimination. *Social Inclusion*, *12*, 1–3. https://doi.org/10.17645/si.8942

Page, C. (2023). Academic language development and linguistic discrimination: Perspectives from internationally educated students. *Comparative and International Education*, *52*(2), 39–53. https://doi.org/10.5206/cie-eci.v52i2.15000

Rogers, S. E. (2016). Bridging the 21st century digital divide. *TechTrends*, *60*(3), 197–199. https://doi.org/10.1007/s11528-016-0057-0

Sohail, A., & Zhang, L. (2024). *Integrating large language models into the psychological sciences*. https://doi.org/10.1007/s12144-025-07438-2

Solovyev, V. D., & Akhtyamova, S. (2019). Linguistic Big Data: Problem of Purity and Representativeness. In *21st International Conference on Data analytics and management in data intensive domains, DAMDID/RCDL 2019* (pp. 193–204). EDN: https://elibrary.ru/tqmgbu

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*. https://doi.org/10.48550/arXiv.1706.03762

Yenduri, G., Ramalingam, M., Chemmalar Selvi, G., Supriya, Y., Srivastava, G., Maddikunta, P. K. R. et al. (2023). Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. In *IEEE Access* (Vol. 12, pp. 54608–54649). https://doi.org/10.1109/access.2024.3389497

Сведения об авторе

Ильин Илья Геннадьевич – магистр права в области информационных технологий, аспирант юридического факультета, Санкт-Петербургский государственный университет

Адрес: 199106, Россия, г. Санкт-Петербург, 22-я линия В.О., 7

E-mail: i.g.ilin@spbu.ru

ORCID ID: https://orcid.org/0000-0003-1076-2765

Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=57765898000

WoS Researcher ID: https://www.webofscience.com/wos/author/record/FDF-0979-2022

Google Scholar ID: https://scholar.google.com/citations?user=YruuMK0AAAAJ

РИНЦ Author ID: https://elibrary.ru/author_profile.asp?authorid=1253542

Конфликт интересов

Автор сообщает об отсутствии конфликта интересов.

Финансирование

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.15 / Конституционное (государственное) право
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 15 ноября 2024 г.
Дата одобрения после рецензирования – 25 ноября 2024 г.
Дата принятия к опубликованию – 25 марта 2025 г.
Дата онлайн-размещения – 30 марта 2025 г.