



IRSTI 04.15.31, 11.07.45

<https://doi.org/10.32523/3080-1702-2026-156-3-63-72>

Scientific article

ALGORITHMIC COMPRESSION OF PUBLIC OPINION: DO LANGUAGE MODELS REPRODUCE THE DIVERSITY OF POLITICAL ATTITUDES?

B. Mendybayev* 

CIOFF Kyrgyzstan, Bishkek, Kyrgyzstan

E-mail: bolashakresearchlab@gmail.com

Abstract. The study draws on two sets of responses to the same political-value questionnaire: answers from 969 participants in Kazakhstan and 300 profiles generated by ChatGPT, Copilot, and Alice. It compares their average scores, the spread of responses, and the distribution of answers across the ten-point scale. The AI-generated data were around 56% less variable than the human responses and covered fewer points on the scale. In other words, the models produced a noticeably narrower range of political positions. This narrowing took two main forms. In some cases, extreme answers were rare, and responses remained close to the middle of the scale. In others, answers clustered near positions that are commonly regarded as socially acceptable.

The extent and form of compression varied across the three AI systems. These findings demonstrate that the similarity between AI-generated responses and the mean position of human respondents does not constitute an adequate representation of public opinion. The utilisation of generative AI in political analysis thus necessitates an assessment of whether such systems can reproduce the pluralism and internal heterogeneity of public attitudes, rather than merely their dominant or average tendencies.

Keywords: generative artificial intelligence, language models, public opinion, political attitudes, algorithmic compression, political pluralism, Kazakhstan.

Introduction

Generative artificial intelligence is becoming part of everyday information, education, and political communication. Systems such as ChatGPT, Copilot, and Alice are now used for much more than searching for information, translating texts, or improving writing. They also explain political events, interpret ideological positions, and provide assessments of institutions, social norms, and public policies. This growing use of conversational AI creates a methodological problem for social and political research. This raises a basic question: do these answers reflect public opinion, or do they mainly reflect how the models were trained and designed?

Existing research demonstrates that machine-generated responses cannot be regarded as neutral reflections of social reality. One of the fundamental limitations of machine-learning systems is their opacity: even when an output is observable, researchers may be unable to reconstruct precisely which training data, optimisation procedures,

Received: 08.08.2026; Revised: 15.08.2026; Approved: 28.09.2026; Available online: 30.09.2026

*the corresponding author

architectural choices, and alignment mechanisms produced it (Burrell, 2016). This limitation is particularly important in the social sciences because political attitudes, institutional trust, views on democracy, and understandings of social justice are not merely informational phenomena. They are formed within specific social, cultural, linguistic, and historical contexts.

Sociologists and political scientists increasingly rely on computational methods to examine large bodies of text. Such methods make it possible to identify recurring themes, trace relationships between concepts, and study public communication across datasets that would be difficult to analyse manually (Evans & Aceves, 2016; Macanovic, 2022; Santurkar *et al.*, 2023). Generative AI presents a rather different challenge. Most earlier tools were designed to analyse material written by people. Large language models also generate the material being studied. Their responses may look like ordinary survey answers, but there is no reason to assume that they reproduce the range, frequency, or internal contradictions of views found among human respondents.

Research based on a quantitative framework found that language models do not represent all social groups equally well (Santurkar *et al.*, 2023). Their responses were closer to the views of some groups than to those of others. For this reason, model outputs cannot be assumed to reflect public opinion. Other researchers found a similar problem. When they compared ChatGPT's answers with the positions of political parties and electoral constituencies, they identified a systematic political bias (Motoki, Pinho Neto & Rodrigues, 2024). Other studies similarly demonstrate that the political orientations of AI systems can be measured, while emphasising that the results depend on the instruments employed, the wording of prompts and survey questions, and the procedures used to interpret model outputs (Rotaru, Anagnoste & Oancea, 2024; Rozado, 2025).

Most existing studies focus primarily on the direction of political bias: whether a model appears more liberal or conservative, closer to a particular party or ideology, or more representative of one demographic group than another. From the perspective of political sociology, however, the location of the mean AI-generated response is only one dimension of representativeness. An equally important question is whether a model can reproduce the internal heterogeneity of public opinion. When a language model systematically reduces variation among responses, it presents a compressed and potentially distorted representation of political space, even when its mean score is close to that of a human sample.

The term algorithmic compression of public opinion describes the loss of variation that occurs when AI-generated answers are compared with responses from actual survey participants. In practice, this may appear as lower variance, repeated use of the same points on a scale, a reduced response range, and fewer extreme answers. The issue is therefore not simply whether an AI system gives a politically moderate or statistically average response. What matters is whether it can reproduce the disagreement, inconsistency, and diversity that are normally present in public opinion.

The importance of this problem is heightened by the growing use of generative AI in political communication, analytical reporting, education, and the preliminary interpretation of public attitudes. When these technologies are incorporated into institutional processes, the resulting outputs can be perceived as objective representations of public opinion. That is not true; automated tools reflect organisational context, technological constraints and value judgments embedded in their design and use (Cole *et al.*, 2022). Within the realm of politics, a language model is not just regurgitating whatever views it has access to. It can cause certain views to seem more rational or prevalent than they truly are and obscure other opinions.

Accordingly, this study examines whether generative language models reproduce the diversity of political attitudes expressed by human respondents or instead exhibit algorithmic compression. The empirical analysis draws on survey data collected from respondents in Kazakhstan and on responses generated by three conversational AI systems to an identical set of political and value-related questions. The principal contribution of the study is to conceptualise generative AI not only as a potentially biased source of political responses but also as a mechanism that may narrow the visible range of public opinion. This shows that a model's representativeness is determined by its ideological alignment as well as its ability to capture the range of viewpoints held by humans.

Literature review

Research on generative artificial intelligence in the social and political sciences has developed at the intersection of several fields, including studies of algorithmic opacity, political bias in AI systems, the simulation of

public opinion through large language models, and the role of digital technologies in specific national and institutional contexts. For this study, the central issue is not only which political positions language models tend to produce, but also whether they can reproduce the internal diversity and distributional complexity of public opinion.

A core issue in this field concerns the challenge of understanding algorithmic systems. Burrell emphasises that machine learning introduces a unique kind of opacity, in which the results produced by an algorithm do not transparently convey the processes behind their formation (Burrell, 2016). Therefore, responses generated by AI should not be regarded as direct or unequivocal representations of social or political perspectives.

Another line of work looks at algorithmic bias. Fazelpour and Danks find that different layers are responsible for bias: the training data, the formulation of the problem, the selection of performance metrics, and the institutional environment that supports the building and use of a given system (Fazelpour & Danks, 2021). Johnson argues that algorithms are not the source of social inequalities and prejudices only when they make technical mistakes, but also because they rely on information shaped by the social system at a given time and thus can mirror these inequalities (Johnson, 2021). Applied to political attitudes, this suggests that the problem is not limited to whether AI systems produce factually incorrect or ideologically biased answers. They may also reproduce a normalised model of political reasoning that privileges some positions, forms of knowledge, and modes of expression over others.

A significant part of the research looks into how large language models relate to politics. Peters describes algorithmic political bias as a consistent pattern where AI systems tend to favour certain political values, groups, or ideologies (Peters, 2022). Motoki and colleagues have found that ChatGPT can show persistent political biases when answering questions about party competition and ideological preferences (Motoki, Pinho Neto and Rodrigues, 2024).

Traditional bias research focuses on whether a model seems more liberal or conservative, institutionalist or anti-establishment, or aligns more with one party, ideology, or social group than another. For example, Rotaru et al. (2024) examined political bias in the way large language models assess media sources and electoral information and found a broadly similar pattern. Similar approaches added a qualification on how judgments about the credibility of news sources may vary with the political orientation associated with those sources (Yang & Menczer, 2025). While this perspective is helpful, it does not fully capture the complexity of how political opinion diversity is represented. Further, an “Ethico-political sensibility” assessment shows that certain specific normative and political biases may be transferred by LLMs (Martin, 2023), and even demonstrate sycophancy effects (Batzner *et al.*, 2024). As a result, congruence between citizens, political institutions, and governing actors raises questions of reliable political representativeness.

Research on ideological and policy congruence emphasises that democratic representation involves far more than simply measuring proximity between the average positions of citizens and political elites (Golder & Stramski, 2010). It also depends on whether institutions respond to the actual distribution of public preferences. Similar research shows that multidimensional congruence between citizens and governments is associated with higher levels of satisfaction with democracy (Stecker & Tausendpfund, 2016). Representation cannot be adequately assessed through a single aggregate indicator (Hooghe, Dassonneville & Oser, 2019), i.e., a model that approximates the average opinion while omitting disagreement, polarisation, or minority positions cannot be considered representative in a politically meaningful sense.

The political attitudes in Kazakhstan are shaped by a complex digital environment. The country has been systematically developing its digital public services, the success of which depends on trust, perceived usefulness, accessibility, and administrative capacity (Kurmanov & Knox, 2024). However, such expansion does not necessarily encourage transparency, inclusion, or civic engagement (Knox & Janenova, 2019), indicating that digitalisation has both administrative and political functions (Keegan, 2024).

Bekmagambetov et al. (2018) investigated the associations between political trust and protest-related behaviour and social media consumption of critical content; Kurmanov and Knox (2024) also discussed the overall technological impact on political trust and citizen-institution relations.

Thus, the literature identifies three points that are central to this study. Artificial intelligence systems can also reflect political and social biases. Second, language models can simulate public opinion, but the performance varies with context, topic, and social group. Third, political attitudes in Kazakhstan and Central Asia are shaped by a unique environment of digital governance, the expansion of social media, and limited political pluralism. What has not been sufficiently explored is whether language models not only reproduce average political positions but

also the underlying diversity of public opinion. This study fills that gap by introducing the concept of algorithmic compression and using it to compare human respondents and generative artificial intelligence systems.

With the provided demographic details and context, it is possible to create “silicon samples” to mimic the responses of human subgroups (Argyle *et al.*, 2023). However, an optimistic view of the possible analytical use of AI responses is not confirmed: data cannot be regarded as a dependable source (Qu & Wang, 2024), and some social groups are more accurately represented than others (Santurkar *et al.*, 2023).

Methods

The methodology utilises a comparative quantitative approach to analyse political and value-orientation statements produced by human respondents and generative AI models. Empirical data include responses to a Kazakhstan-wide survey and responses of three chatbots to the same survey questionnaire. The study consisted of 12 items rated on ten-point scales. Questions related to the value of free elections and democracy, faith in strong political leadership, the legitimacy of income inequality, belief that it is the state’s duty to take care of its citizens, public order or high-value freedoms, justifiability of disobeying an unfair law, importance of traditions, national identity, and confidence in scientists, government, and media.

The analysis encompassed differences in means as well as the distributional features of responses. A comparison was made between human and AI-sampled respondents based on the dispersion and concentration of their answers and the number of categories utilised. The examination assessed whether language models include the full range of answers or limit their output to a few categories. “Algorithmic compression of public opinion” happens if AI answers are more limited in range than survey responses. This compression does not refer to a change in the substantive meaning of an answer but to a loss of variation—responses occupying fewer points on the scale, clustering more tightly around certain values, or exhibiting less dispersion. The term also does not imply that an AI system holds a political opinion; rather, the analysis considers the pattern produced when the system responds to questions about political values.

The empirical comparison combines a human survey with 300 AI-generated questionnaires. The survey was conducted in Kazakhstan and produced 1,015 completed questionnaires. After data-quality checks, 969 cases were retained. Participation was voluntary and anonymous. The sample includes a comparatively large proportion of young, urban, and student respondents and is not nationally representative. Its purpose in this study is more limited. It shows how answers to the 12 items were distributed within an observed group of respondents and provides a human distribution against which the AI results can be compared.

Compression was evaluated based on various features of the resulting distributions, including their dispersion, the proportion of responses at mode values, and the number of scale categories utilised. Normalised Shannon entropy was included because the number of categories alone does not show whether responses are distributed evenly across them. Entropy was calculated for each item and then averaged across the 12-item battery. Higher values indicate that answers are spread more evenly across the ten available categories. Lower values indicate that a relatively small part of the scale accounts for most responses.

The analysis was conducted first with all 300 AI questionnaires combined and then repeated separately for ChatGPT, Copilot, and Alice. The first comparison shows whether a narrower response distribution is visible across the AI data. The separate comparisons indicate whether the same pattern is produced by all three systems or is driven mainly by one of them. AI-generated responses were interpreted not as expressions of autonomous political opinion, but as algorithmically produced response patterns to questions concerning political values. This analytical distinction allows the study to assess the extent to which such systems reproduce the pluralism and distributional heterogeneity of public opinion in Kazakhstan.

Results

The main difference between the human and AI samples was not limited to the mean position on individual questions. AI-generated answers were also distributed across a narrower part of the scale. For each of the 12 items, the standard deviation was lower in the pooled AI sample than among human respondents. Averaged across the questionnaire, the AI-to-human standard deviation ratio was 0.435. AI responses therefore showed approximately 56% less variation than the survey responses.

Other distributional measures point in the same direction. As shown in Table 1, the most frequent response category accounted, on average, for 49.6% of the pooled AI responses, compared with 27.5% in the human sample. Normalised Shannon entropy was also markedly lower for AI-generated data: 0.547 against 0.887. Human respondents were thus more evenly distributed across the available response categories, whereas AI answers were concentrated around a smaller set of values.

Table 1. Overall indicators of the diversity of human and AI responses

Indicator	Human respondents	AI systems	Interpretation
Mean share of the modal response	27.5%	49.6%	Nearly half of all AI responses fall within the most frequent category.
Mean normalised entropy	0.887	0.547	AI-generated distributions are substantially less diverse.
Mean AI-to-human standard deviation ratio	—	0.435	The dispersion of AI responses is considerably lower.
Mean share of extreme responses	33.6%	23.2%	AI systems use the endpoints of the scale less frequently.
Mean share of middle-range responses	48.4%	55.6%	AI responses are more concentrated in the moderate part of the scale.

Note: Compiled by the author based on the survey and AI-generated questionnaire data.

The difference was also visible in the use of the scale. Extreme responses represented 33.6% of the human answers and 23.2% of the AI answers. Conversely, middle-range values accounted for 55.6% of AI responses, compared with 48.4% among human respondents. The models did not simply reproduce the human distribution with different mean scores. They reduced their spread and assigned greater weight to a limited part of the response scale.

This narrowing did not take the same form across all questions. For the institutional-trust items, AI responses were concentrated mainly in the middle of the scale. A similar tendency appeared in questions concerning social responsibility, traditions, and national identity. Here, the principal difference from the human sample was the reduced presence of responses at the extremes.

Free elections, democracy, and strong leadership produced a different type of narrowing. The AI systems repeatedly combined high ratings of elections and democracy with low ratings of rule by a strong leader unconstrained by parliament or elections. Their answers did not cluster around the middle of the scale. Instead, they reproduced a relatively fixed procedural-democratic profile, leaving little room for the combination of democratic support and acceptance of strong leadership found among human respondents.

The pooled AI results conceal substantial differences between the three systems. Copilot generated the broadest distributions. It used an average of 5.42 scale categories per item, its modal response accounted for 43.2% of answers, and its standard deviation was 49.3% of the corresponding human value.

ChatGPT produced more varied responses than Alice, although both systems used the scale more narrowly than the human respondents. ChatGPT selected an average of 3.75 categories per item, but 66.5% of its answers still fell within the modal category. Its standard deviation was 37.4% of the corresponding value in the human sample. Alice's responses were considerably more uniform. It used an average of 2.42 categories per item, and 91.3% of its answers coincided with the modal value. Its standard deviation was only 13.5% of that observed among human respondents. Repeated administrations therefore generated some variation in ChatGPT's answers, but very little in Alice's, as shown in Table 2.

Table 2. Differences between AI systems in terms of the degree of compression of responses

AI system	Mean AI-to-human standard deviation ratio	Mean number of scale categories used	Mean share of the modal response
ChatGPT	0.374	3.75	66.5%
Copilot	0.493	5.42	43.2%
Alice	0.135	2.42	91.3%

Note: Compiled by the author based on the survey and AI-generated questionnaire data.

The results identify empirically distinct forms of compression. In the first, answers move toward the middle of the scale and extreme positions become less frequent; this reflects intensity. The algorithmic compression of public opinion thus becomes evident: responses tend to cluster around a single pole, as observed in the items concerning democracy, elections, and strong leadership, thereby narrowing the spectrum of positions around a preferred orientation. Consequently, both patterns serve to diminish distributional diversity, albeit through different mechanisms.

Discussion

The comparison between human and AI responses changes when shifting focus from the mean to the distribution used for calculation. Human respondents combined moderate and common positions with strongly expressed, infrequent, and mixed answers that formed part of the political-value profile. The three AI systems depended heavily on a limited set of recurring values, which caused similar averages to obscure substantial differences in variation. Algorithmic compression refers to this process of narrowing the response distribution, and its specific form is important.

Response patterns that cluster around a single normative point can create the illusion of uncontested consensus and limit the range of acceptable alternatives. Increased consistency in responses does not necessarily serve as evidence of enhanced representational accuracy. Political attitudes are seldom organised as a coherently articulated ideology, particularly in contexts where party competition does not offer citizens stable and comprehensive packages of political positions. Naturally, human respondents may provide contradictory positions when mapped onto conventional ideological scales, as this reflects diverse conceptualisations of authority, order, justice, and responsibility.

Analytical models alter the underlying structure of political meaning by excluding complex combinations in the absence of validation. This issue is particularly relevant to the Kazakhstan case. The human survey profile does not fit a single liberal, conservative, democratic, or authoritarian position. It contains several value configurations that have developed within a specific institutional and historical setting. The models arrange these orientations more sharply, especially where democratic procedures and strong authority are concerned. Their answers may therefore be coherent in general normative terms while remaining distant from the way these values coexist among the surveyed respondents. The ability to communicate in Russian or serve users in the region does not mean that a system has acquired a locally grounded political-value profile.

ChatGPT, Copilot, and Alice prompt testing illustrated that models' answers vary even under the same questions and conditions, making it hard to define a single AI view of public opinion. While the comparison does not identify the process behind these differences, it highlights that conclusions from one model cannot be applied to others. Algorithmic compression is distinct from political bias—bias shows where a model sits in political space, while compression shows how much of that space is reflected in its responses. A model might align with the human average but omit disagreement or favour a certain normative stance with varying expression. Political analysis needs both to assess if diversity is preserved, not just the direction.

Thus, the apparent similarity between the average positions of human respondents and those generated by language models should not be taken as evidence of adequate representativeness of public opinion. Algorithmic compression reduces statistical variation and obscures the contradictions and pluralism that constitute the very essence of political attitudes—such outputs risk creating the illusion of greater coherence and moderation than is present in society.

Conclusions

AI-generated responses were less dispersed than human responses on all 12 questionnaire items: the models repeatedly concentrated their answers within a narrower set of positions. Two forms of algorithmic compression were identified. The first, middle-range compression, is characterised by the concentration of responses in the moderate part of the scale and the underrepresentation of extreme positions. The second, normative-pole compression, occurs when responses cluster around a socially approved end of the scale, such as support for democracy and free elections or rejection of unconstrained personal rule. Algorithmic compression should therefore not be understood simply as a tendency towards moderation. Its defining feature is the reduction in the range and diversity of positions represented.

The degree of compression varied across the three AI systems, thus confirming that the representation of political attitudes depends on model-specific characteristics and on the procedures through which responses are generated. A language model may produce responses that are close to the average position of a human sample while still omitting real disagreement, polarisation, minority views, and internally contradictory attitudes. In such cases, apparent similarity conceals a simplified and less pluralistic representation of public opinion.

Generative AI should not be seen as an independent source of public opinion, as dispersion, concentration, and pluralism must be preserved. For Kazakhstan, where digital media and algorithms increasingly influence politics and socialisation, AI in public opinion research and communication might unintentionally depict society as more homogeneous than it is. Therefore, AI should be a supplementary tool for analysis and hypothesis generation, not a standalone source of data.

Ethics Statement

The study was reviewed and approved by the Research Ethics Committee of L.N. Gumilyov Eurasian National University (Protocol No. 2, dated 29 September 2025). Participation was voluntary and based on informed consent. The survey was conducted anonymously, without collecting names, contact details, personal identification numbers, or other directly identifying information. Participants were informed about the purpose of the study, their right to withdraw at any stage, and the use of the results exclusively in aggregated form.

Funding

This work was supported by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan under Grant AP26104035.

Data Availability

The dataset, questionnaire, research protocol, data dictionary, coding materials, and aggregated tables supporting the findings of this study are available in the Zenodo repository at <https://doi.org/10.5281/zenodo.20339263>.

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Argyle, L.P., Busby, E.C., Fulda, N., Gubler, J.R., Rytting, C. and Wingate, D. (2023) 'Out of one, many: Using language models to simulate human samples', *Political Analysis*, 31(3), pp. 337–351. Available at: <https://doi.org/10.1017/pan.2023.2>.
2. Batzner, J., Stocker, V., Schmid, S. and Kasneci, G. (2024) 'GermanPartiesQA: Benchmarking commercial large language models and AI companions for political alignment and sycophancy'. *arXiv*. Available at: <https://doi.org/10.48550/ARXIV.2407.18008>.
3. Bekmagambetov, A., Wagner, K.M., Gainous, J., Sabitov, Z., Rodionov, A. and Gabdulina, B. (2018) 'Critical social media information flows: Political trust and protest behaviour among Kazakhstani college students', *Central Asian Survey*, 37(4), pp. 526–545. Available at: <https://doi.org/10.1080/02634937.2018.1479374>.
4. Burrell, J. (2016) 'How the machine "thinks": Understanding opacity in machine learning algorithms', *Big Data & Society*, 3(1), p. 2053951715622512. Available at: <https://doi.org/10.1177/2053951715622512/>

5. Cole, M., Cant, C., Ustek Spilda, F. and Graham, M. (2022) 'Politics by automatic means? A critique of artificial intelligence ethics at work', *Frontiers in Artificial Intelligence*, 5, p. 869114. Available at: <https://doi.org/10.3389/frai.2022.869114>.
6. Evans, J.A. and Aceves, P. (2016) 'Machine translation: Mining text for social theory', *Annual Review of Sociology*, 42(1), pp. 21–50. Available at: <https://doi.org/10.1146/annurev-soc-081715-074206>.
7. Fazelpour, S. and Danks, D. (2021) 'Algorithmic bias: Senses, sources, solutions', *Philosophy Compass*, 16(8), p. e12760. Available at: <https://doi.org/10.1111/phc3.12760>.
8. Golder, M. and Stramski, J. (2010) 'Ideological congruence and electoral institutions', *American Journal of Political Science*, 54(1), pp. 90–106. Available at: <https://doi.org/10.1111/j.1540-5907.2009.00420.x>.
9. Hooghe, M., Dassonneville, R. and Oser, J. (2019) 'Public opinion, turnout and social policy: A comparative analysis of policy congruence in European liberal democracies', *Political Studies*, 67(4), pp. 992–1009. Available at: <https://doi.org/10.1177/0032321718819077>.
10. Johnson, G.M. (2021) 'Algorithmic bias: on the implicit biases of social technology', *Synthese*, 198(10), pp. 9941–9961. Available at: <https://doi.org/10.1007/s11229-020-02696-y>.
11. Keegan, K. (2024) 'Corruption and digital authoritarianism: political drivers of e-government adoption in Central Asia', *Democratization*, 31(1), pp. 23–46. Available at: <https://doi.org/10.1080/13510347.2023.2255146>.
12. Knox, C. and Janenova, S. (2019) 'The e-government paradox in post-Soviet countries', *International Journal of Public Sector Management*, 32(6), pp. 600–615. Available at: <https://doi.org/10.1108/IJPSM-08-2018-0173>.
13. Kurmanov, B. and Knox, C. (2024) 'Digital activism and authoritarian legitimization in post-Soviet Central Asia', *The Information Society*, 40(5), pp. 343–360. Available at: <https://doi.org/10.1080/01972243.2024.2374714>.
14. Macanovic, A. (2022) 'Text mining for social science – The state and the future of computational text analysis in sociology', *Social Science Research*, 108, p. 102784. Available at: <https://doi.org/10.1016/j.ssresearch.2022.102784>.
15. Motoki, F., Pinho Neto, V. and Rodrigues, V. (2024) 'More human than human: measuring ChatGPT political bias', *Public Choice*, 198(1–2), pp. 3–23. Available at: <https://doi.org/10.1007/s11127-023-01097-2>.
16. Peters, U. (2022) 'Algorithmic political bias in artificial intelligence systems', *Philosophy & Technology*, 35(2), p. 25. Available at: <https://doi.org/10.1007/s13347-022-00512-8>.
17. Qu, Y. and Wang, J. (2024) 'Performance and biases of large language models in public opinion simulation', *Humanities and Social Sciences Communications*, 11(1), p. 1095. Available at: <https://doi.org/10.1057/s41599-024-03609-x>.
18. Rotaru, G.-C., Anagnoste, S. and Oancea, V.-M. (2024) 'How artificial intelligence can influence elections: Analyzing the Large Language Models (LLMs) Political Bias', *Proceedings of the International Conference on Business Excellence*, 18(1), pp. 1882–1891. Available at: <https://doi.org/10.2478/picbe-2024-0158>.
19. Rozado, D. (2025) 'Measuring political preferences in AI Systems: An integrative approach'. *arXiv*. Available at: <https://doi.org/10.48550/ARXIV.2503.10649>.
20. Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P. and Hashimoto, T. (2023) 'Whose opinions do language models reflect?', *Proceedings of the 40th International Conference on Machine Learning, Proceedings of Machine Learning Research*, 202, pp. 29971–30004.
21. Stecker, C. and Tausendpfund, M. (2016) 'Multidimensional government-citizen congruence and satisfaction with democracy', *European Journal of Political Research*, 55(3), pp. 492–511. Available at: <https://doi.org/10.1111/1475-6765.12147>.
22. Yang, K.-C. and Menczer, F. (2025) 'Accuracy and political bias of news source credibility ratings by large language models', *Proceedings of the 17th ACM Web Science Conference 2025. Websci '25: 17th ACM Web Science Conference 2025, New Brunswick NJ USA: ACM*, pp. 127–137. Available at: <https://doi.org/10.1145/3717867.3717903>.

Б. Меңдібаев

CIOFF Kyrgyzstan, Бишкек, Кыргызстан

ҚОҒАМДЫҚ ПІКІРДІ АЛГОРИТМДІК СЫҒЫМДАУ: ТІЛДІК МОДЕЛЬ- ДЕР САЯСИ ҰСТАНЫМДАРДЫҢ ӘРТҮРЛІЛІГІН ҰДАЙЫ ӨНДІРЕ МЕ?

Аңдатпа. Мақалада генеративті жасанды интеллект жүйелерінің саяси ұстанымдардың әртүрлілігін жаңғырту қабілеті қарастырылады. Зерттеудің эмпирикалық негізін Қазақстандағы 969 респонденттің жауаптары және ChatGPT, Copilot пен Алиса жүйелері бірдей сұрақтар бойынша қалыптастырған 300 сауалнамалық профиль құрады. Әдіснамалық тұрғыдан зерттеу жауаптардың орташа мәндерін, стандартты ауытқуын, модальды және шеткі жауаптардың үлесін, пайдаланылған шкала санаттарының санын және нормаланған Шеннон энтропиясын салыстыруға негізделді.

Талдау барысында орташа бағалармен қатар он балдық шәкілдегі жауаптардың шашырауы, шоғырлануы және әртүрлі позициялардың ұсынылуы салыстырылды. ЖИ жауаптарының өзгергіштігі адамдар жауаптарының өзгергіштігіне қарағанда орта есеппен 56%-ға төмен екені анықталды. Тілдік модельдердің жауаптары жекелеген мәндердің төңірегіне жиі шоғырланып, саяси ұстанымдардың тар ауқымын қамтиды. Алгоритмдік сығымдаудың екі түрі анықталды: шеткі бағалардың үлесі төмендейтін ортаңғы сығымдау және жауаптар әлеуметтік тұрғыдан құпталатын позицияға шоғырланатын нормативтік-полюстік сығымдау. ЖИ жүйесінің респонденттердің орташа позициясына жақындығы қоғамдық пікірді толық жаңғыртуды білдірмейді.

Зерттеудің ғылыми үлесі алгоритмдік сығымдауды саяси бейімділіктен бөлек, жасанды интеллект қалыптастыратын жауаптардың репрезентативтілігін бағалаудың дербес өлшемі ретінде ұсырудан тұрады. Саяси талдауда генеративті ЖИ-ді пайдалану кезінде модельдердің басым ұстанымдармен қатар қоғамдағы пікірлер алуандығын көрсету қабілетін де ескеру қажет.

Негізгі ұғымдар: генеративті жасанды интеллект, тілдік модельдер, қоғамдық пікір, саяси ұстанымдар, алгоритмдік сығымдау, саяси плюрализм, Қазақстан.

Б. Мендыбаев

CIOFF Kyrgyzstan, Бишкек, Кыргызстан

АЛГОРИТМИЧЕСКОЕ СЖАТИЕ ОБЩЕСТВЕННОГО МНЕНИЯ: ВОСПРОИЗВОДЯТ ЛИ ЯЗЫКОВЫЕ МОДЕЛИ РАЗНООБРАЗИЕ ПОЛИТИЧЕСКИХ УСТАНОВОК?

Аннотация. В статье рассматривается способность генеративных систем искусственного интеллекта воспроизводить разнообразие политических установок. Исследовательская проблема связана с тем, что близость ответов языковых моделей к средним значениям человеческой выборки не позволяет определить, насколько полно они отражают внутреннюю неоднородность общественного мнения. Эмпирическую основу исследования составили ответы 969 респондентов из Казахстана и 300 анкетных профилей, сформированных ChatGPT, Copilot и Алисой по идентичным вопросам.

Методология основана на сравнительном анализе распределений ответов: сопоставлялись средние значения, стандартные отклонения, доля модальных и крайних ответов, число используемых категорий шкалы и нормированная энтропия Шеннона. Установлено, что вариативность ответов ИИ в среднем на 56% ниже вариативности человеческих ответов, при этом степень сжатия различается между системами. Выявлены две формы алгоритмического сжатия: срединная и нормативно-полюсная.

Научный вклад исследования состоит в концептуализации алгоритмического сжатия общественно-го мнения как отдельного измерения репрезентативности ИИ, отличного от политической предвзятости. Результаты показывают, что оценка применения генеративного ИИ в политическом анализе должна учитывать не только близость к средним установкам, но и способность моделей сохранять разнообразие, противоречивость и плюрализм общественного мнения.

Ключевые слова: генеративный искусственный интеллект, языковые модели, общественное мнение, политические установки, алгоритмическое сжатие, политический плюрализм, Казахстан.

Information about the author

B. Mendybayev – PhD, researcher in sociology and political studies, CIOFF Kyrgyzstan. Bishkek, Kyrgyzstan

Автор туралы мәлімет

Б. Меңдібаев – PhD, әлеуметтану және саяси зерттеулер саласындағы зерттеуші, CIOFF Kyrgyzstan. Бішкек, Қырғызстан

Сведения об авторе

Б. Мендыбаев – PhD, исследователь в области социологии и политических исследований, CIOFF Kyrgyzstan, Бишкек, Кыргызская Республика