УДК 316.77:004.77

Гэ Аньгэ
Магистр
Казахский национальный университет имени аль-Фараби
Алматы, Казахстан

ЭМПИРИЧЕСКОЕ ИССЛЕДОВАНИЕ ФОРМИРОВАНИЯ И ГРУППОВЫХ РАЗЛИЧИЙ АЛГОРИТМИЧЕСКИХ КОКОНОВ РОЛЬ ПРОФИЛИРОВАНИЯ ПОЛЬЗОВАТЕЛЕЙ В АЛГОРИТМИЧЕСКИХ РЕКОМЕНДАЦИЯХ С ПОЗИЦИЙ ТЕОРИИ ЦИФРОВОГО РАЗРЫВА И ТЕОРИИ УСТАНОВЛЕНИЯ ПОВЕСТКИ ДНЯ

Аннотация. В статье рассматривается формирование алгоритмических коконов и их групповые различия среди интернет-пользователей Казахстана. Актуальность исследования определяется возрастающей ролью цифровых платформ и рекомендательных алгоритмов в структурировании информационной среды. Цель работы состоит в выявлении влияния структурных условий, цифровых компетенций и поведенческих практик пользователей на степень замкнутости алгоритмического кокона и воспринимаемое разнообразие повестки. Теоретическую основу исследования составляют теория цифрового разрыва и теория установления повестки дня. Эмпирическая база включает данные опроса 200 пользователей интернета в Казахстане. В исследовании применяются корреляционный анализ, групповые сравнения, иерархическая регрессия и последовательный медиативный анализ. Результаты показывают, что образование, место проживания, профессиональный статус и языковая среда существенно влияют на цифровые компетенции и информационное поведение пользователей. Более высокие цифровые компетенции способствуют активному поиску информации, снижению замкнутости алгоритмического кокона и росту воспринимаемого разнообразия повестки.

Ключевые слова: Алгоритмический Кокон, Цифровой Разрыв, Рекомендательные Алгоритмы, Информационное Неравенство, Казахстан.

Ge Ange
Master
Al-Farabi Kazakh National University
Almaty, Kazakhstan

AN EMPIRICAL STUDY ON THE FORMATION AND GROUP DIFFERENCES OF ALGORITHMIC COCOONS THE ROLE OF USER PROFILING IN ALGORITHMIC RECOMMENDATION FROM THE PERSPECTIVES OF DIGITAL DIVIDE AND AGENDA-SETTING THEORIES

Abstract. The article examines the formation of algorithmic cocoons and their group differences among internet users in Kazakhstan. The relevance of the study is determined by the growing role of digital platforms and recommendation algorithms in structuring the information environment. The aim of the research is to identify the influence of structural conditions, digital competencies, and users' behavioural practices on the tightness of algorithmic cocoons and perceived agenda diversity. The theoretical basis of the study is formed by digital divide theory and agenda-setting theory. The empirical basis includes survey data from 200 internet users in Kazakhstan. The study applies correlation analysis, group comparison, hierarchical regression, and serial mediation analysis. The results show that education, place of residence, occupational status, and language environment significantly influence users' digital competencies and information behaviour. Higher digital competencies contribute to active information seeking, reduce algorithmic cocoon tightness, and increase perceived agenda diversity.

Keywords: Algorithmic Cocoon, Digital Divide, Recommendation Algorithms, Information Inequality, Kazakhstan.

Introduction

The rapid development of digital platforms and algorithmic recommendation systems has significantly changed the mechanisms of information dissemination.

Social media platforms increasingly determine not only the speed and scale of information circulation, but also the structure of users' everyday informational environments. By collecting and analysing behavioural data, recommendation algorithms construct user profiles and provide individualized content streams. Although personalization increases the efficiency of information delivery, it may also narrow informational diversity, reinforce homogeneous exposure, and contribute to the formation of algorithmic cocoons.

The relevance of this problem is connected with the growing dependence of citizens on digital platforms as sources of news, social communication, and public discussion. In algorithmically mediated environments, access to diverse information is shaped not only by individual preferences or media institutions, but also by platform infrastructures, recommendation mechanisms, and users' digital capacities. As a result, algorithmic personalization may reproduce and intensify existing social inequalities, leading to unequal informational outcomes among different social groups.

Existing studies on filter bubbles, echo chambers, and algorithmic cocoons have contributed to the understanding of informational closure. However, much of this research remains focused on Western media systems and pays insufficient attention to non-Western, transitional, and multilingual contexts. The relationship between structural inequality and informational inequality in algorithmic environments also remains insufficiently explored. In particular, it is important to examine how education, residence, occupation, language environment, and digital literacy influence algorithmic cocoon tightness and perceived agenda diversity.

Kazakhstan represents a relevant case for such analysis. The country is characterized by rapid digital transformation, active use of Instagram, TikTok, Telegram, and YouTube, persistent urban–rural and educational inequalities, and a bilingual information environment formed by Kazakh- and Russian-language media spaces. These conditions make it possible to examine how algorithmic recommendation systems may reinforce both social and linguistic segmentation of information flows.

The purpose of this study is to analyse the formation of algorithmic cocoons and group differences in their manifestation among internet users in Kazakhstan. The study examines how structural conditions, digital capacities, and behavioural practices influence algorithmic cocoon tightness and perceived agenda diversity.

The theoretical basis of the research is formed by digital divide theory and agenda-setting theory. Digital divide theory explains how social inequalities are transformed into differences in digital capacities and platform behaviour. Agenda-setting theory makes it possible to analyse how algorithmic systems shape users' exposure to different issues, topics, and viewpoints. The integration of these approaches allows algorithmic inequality to be considered as a multidimensional phenomenon involving structural conditions, digital skills, user practices, and informational outcomes.

The empirical basis of the study consists of survey data collected from 200 internet users in Kazakhstan, supplemented by qualitative interpretation. The analysis includes descriptive statistics, reliability and validity testing, correlation analysis, group comparison, hierarchical regression, and serial mediation analysis.

The scientific novelty of the study lies in the integration of digital divide and agenda-setting perspectives for the analysis of algorithmic cocoons in a multilingual Central Asian context. Unlike approaches that interpret algorithmic cocoons mainly as technological effects, this study considers them as socially embedded phenomena shaped by structural inequality, digital capacities, and user behaviour. The practical significance of the research is related to the development of more inclusive approaches to digital literacy, platform governance, and informational diversity.

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

The development of social media platforms and algorithmic recommendation systems has significantly transformed the logic of information dissemination. Communication is increasingly shifting from a mass-media model toward a platform-based and personalized system of content distribution. By analysing user behaviour, algorithms filter, rank, and recommend information according to presumed preferences. While this increases the efficiency of information delivery, it may also

narrow users' informational horizons and reinforce homogeneous exposure.

In the literature, three related concepts are used to explain this phenomenon: filter bubbles, echo chambers, and algorithmic cocoons. The concept of the filter bubble emphasizes the technological logic of algorithmic filtering. Platforms collect data on users' browsing, searching, and interaction patterns and provide content corresponding to their previous interests, thereby limiting exposure to alternative viewpoints [1]. This perspective highlights the non-neutral character of algorithms and their role in reinforcing existing preferences.

The concept of the echo chamber, in contrast, focuses on user agency and social interaction. Users actively select sources, communities, and networks that correspond to their beliefs, while practices such as following, liking, commenting, and sharing strengthen homogeneous communicative environments [2]. Thus, echo chambers emerge from the interaction between algorithmic recommendation mechanisms and users' selective behaviour.

Recent studies show that users are not entirely passive recipients of algorithmic influence. Through active searching, cross-platform media use, and source verification, some users can reduce the effects of personalization and access more diverse perspectives [6]. However, these capacities are unevenly distributed across social groups, which leads to differentiated algorithmic outcomes.

The concept of the algorithmic cocoon provides a broader analytical framework, as it combines algorithmic mechanisms, user behaviour, and social structure. Unlike filter bubbles and echo chambers, it allows informational closure to be understood not as an absolute state, but as a continuous and measurable condition. The tightness of an algorithmic cocoon may vary depending on education, digital literacy, social position, language environment, and media consumption practices.

From this perspective, algorithmic cocoons can be regarded as a manifestation of social stratification in the digital environment. Users with higher education and stronger digital competence are more capable of understanding recommendation mechanisms, adjusting platform settings, diversifying information sources, and maintaining open informational environments. Conversely, disadvantaged users tend

to rely on default recommendations and are more likely to remain within closed and repetitive information flows. In multilingual contexts such as Kazakhstan, language also structures access to information: Russian-speaking users often enter broader transnational information networks, whereas Kazakh-speaking users are more likely to encounter localized content streams.

Digital divide theory provides an important basis for analysing these inequalities. Early studies focused mainly on unequal access to internet infrastructure [3], while later research emphasized differences in digital skills, usage patterns, and the ability to benefit from online resources. These differences are shaped by education, income, occupation, age, and urban–rural residence. According to van Dijk, digital inequality originates in broader disparities in social resources and is reproduced in digital contexts, where socially advantaged groups develop stronger informational and operational capacities [4].

In the algorithmic era, the digital divide develops into algorithmic inequality. Although recommendation systems are formally accessible to all users, their effects depend on users' digital capacities and behavioural strategies. Users with higher digital competence are more likely to understand algorithmic logic, actively modify preferences, use multiple platforms, and seek heterogeneous information. These practices reduce cocoon tightness. By contrast, users with lower digital competence are more dependent on default feeds and are more likely to consume homogeneous content.

Algorithmic inequality therefore appears in differentiated informational outcomes. Users with different digital capacities encounter different information structures, issue agendas, and levels of exposure to diverse viewpoints. Digital capacities, including algorithmic awareness, media literacy, information search skills, and source verification abilities, mediate the relationship between structural conditions and informational diversity. Behavioural practices such as active searching, cross-platform use, and preference adjustment further determine the openness of users' information environments.

Kazakhstan represents a relevant context for examining these processes. Persistent

urban–rural, educational, and occupational inequalities continue to shape access to infrastructure, digital skills, and information resources. Urban users generally benefit from better technological conditions and higher digital literacy, while rural users face both technical and competence-related limitations. The bilingual information environment further deepens these differences. Russian-speaking users often access wider post-Soviet and transnational media networks, whereas Kazakh-speaking users are more embedded in localized information spaces. Thus, algorithmic cocoons may be understood as an extension of the digital divide in the algorithmic era.

Agenda-setting theory offers another important perspective. Classical agenda-setting theory argues that the media influence public perceptions of issue importance by determining which topics receive attention and visibility [5]. In traditional mass media systems, this process was largely organized by centralized media institutions. In algorithmic environments, however, agenda formation becomes decentralized, personalized, and fragmented. Algorithms generate individualized content feeds based on user data, replacing traditional editorial gatekeeping with dynamic and data-driven filtering mechanisms.

Algorithmic agenda-setting can be described through three features. First, it is decentralized, since issue visibility results from the interaction of algorithms, users, and content circulation. Second, it is individualized, because personalized feeds expose different users to different topics. Third, it is structurally differentiated, as users' social positions and digital capacities influence the diversity of information they encounter.

Perceived agenda diversity is therefore an important indicator of informational openness. It reflects users' subjective assessment of the range of issues, viewpoints, and social perspectives available to them. A tighter algorithmic cocoon is associated with a narrower perceived agenda, while a more open information environment corresponds to higher perceived diversity. In multilingual societies such as Kazakhstan, language further shapes agenda exposure: Russian-speaking users may be more connected to transnational issues, while Kazakh-speaking users may be more oriented toward local and national topics.

Despite the relevance of these theoretical approaches, several gaps remain in the existing literature. First, empirical research on algorithmic communication in Central Asia remains limited. Second, digital divide theory and agenda-setting theory are rarely integrated in the study of algorithmic information environments. Third, multilingual dynamics are often overlooked, although language plays an important role in structuring access to information. Fourth, previous studies have paid insufficient attention to informational outcomes, particularly algorithmic cocoon tightness and perceived agenda diversity.

By focusing on Kazakhstan, this study addresses these gaps and develops a framework explaining how structural conditions, digital capacities, and behavioural practices jointly shape algorithmic cocoon tightness and perceived agenda diversity. The study therefore interprets algorithmic cocoons not merely as technological effects, but as socially embedded phenomena that reproduce and intensify informational inequality in the digital environment.

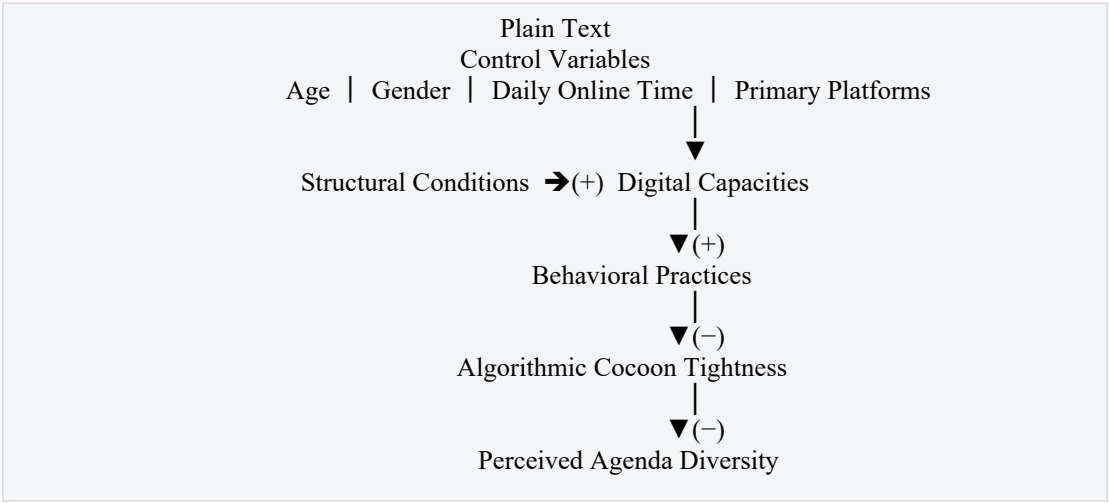

Figure 1 Theoretical Research Framework
Notes: (+) indicates positive effect; (−) indicates negative effect.

This study adopts a sequential explanatory mixed-methods design, with quantitative analysis as the primary approach and qualitative data used for interpretation.

The sample consists of Kazakhstani internet users aged 18–45, selected through stratified sampling. Data were collected via anonymous online questionnaires in Kazakh, Russian, and English. Of 223 responses, 200 valid cases were retained.

All variables were measured using five-point Likert scales, and the

questionnaire was pretested to ensure clarity and validity.

Quantitative analysis was conducted using SPSS 26.0, including reliability testing, descriptive statistics, Pearson correlation, independent-samples t-tests, one-way ANOVA, hierarchical regression, and PROCESS-based mediation analysis to examine chained effects.

A subgroup (robustness) analysis was further conducted by comparing a high-vulnerability group (rural, lower education, lower digital capacity) with an urban high-capacity group, to assess differences in algorithmic cocoon formation and informational outcomes.

All constructs were measured using multiple indicators:

Structural Conditions.

Education (four levels), residence (urban/rural), occupation (five categories), and socioeconomic status (subjective self-assessment).

Digital Capacities.

Information searching ability, source verification, platform operation skills, and algorithmic understanding.

Behavioral Practices.

Proactive searching for diverse viewpoints, cross-platform information use, engagement with heterogeneous content, preference adjustment, and verification behavior.

Algorithmic Cocoon Tightness.

Perceived content repetition, alignment with prior interests, limited exposure to diverse viewpoints, and overall information homogeneity.

Perceived Agenda Diversity.

Perceived exposure to diverse perspectives, range of social issues, and awareness of viewpoints across different groups.

Table 1 presents the demographic profile of the sample. The final dataset includes 200 valid responses. The sample is balanced in gender (47.5% male, 52.5% female), with 63% urban residents and 58.5% holding a bachelor’s degree or above. Average daily online time is 5.7 hours, indicating high platform engagement.

Instagram (82.5%) and Telegram (79.0%) are the dominant platforms.

Table 1 Demographic Characteristics (N=200)

Variable	Category	Frequency	Percentage (%)
Gender	Male	95	47.5
	Female	105	52.5
Age	18–25	82	41.0
	26–35	77	38.5
	36–45	41	20.5
Residence	Urban	126	63.0
	Rural	74	37.0
Education	High school or below	35	17.5
	Vocational college	48	24.0
	Bachelor's or above	117	58.5
Occupation	Student	55	27.5
	Civil servant	33	16.5
	Corporate employee	62	31.0
	Self-employed	24	12.0
	Manual labor	26	13.0

Table 2 reports descriptive statistics of core variables. Algorithmic cocoon tightness shows a relatively higher mean (M = 3.41, SD = 0.65) compared to perceived agenda diversity (M = 3.26, SD = 0.71).

Table 2 Descriptive Statistics of Core Variables (N=200)

Variable	Mean	Standard Deviation
Digital Capacities	3.68	0.62
Behavioral Practices	3.54	0.69
Algorithmic Cocoon Tightness	3.41	0.65
Perceived Agenda Diversity	3.26	0.71

Table 3 Reliability (Cronbach's α)

Variable	Items	Cronbach's α
Digital Capacities	8	0.871
Behavioral Practices	7	0.842
Algorithmic Cocoon Tightness	6	0.826
Perceived Agenda Diversity	7	0.854

Table 4 KMO and Bartlett's Test

Variable	KMO	Bartlett's χ^2	df	p
Digital Capacities	0.882	426.78	28	<0.001
Behavioral Practices	0.846	357.21	21	<0.001
Algorithmic Cocoon Tightness	0.811	298.54	15	<0.001
Perceived Agenda Diversity	0.859	389.67	21	<0.001

All constructs demonstrate good internal consistency, with Cronbach's α ranging from 0.826 to 0.871 (Table 3). KMO values exceed 0.80 and Bartlett's tests are significant at $p < 0.001$ (Table 4), confirming sampling adequacy and construct validity.

Harman's single-factor test indicates that the first factor explains 28.7% of variance, below the 40% threshold, suggesting no serious common method bias.

Table 5 reports Pearson correlations among key variables. Structural conditions are significantly associated with digital capacities ($r = 0.52$, $p < 0.001$) and behavioral practices ($r = 0.48$, $p < 0.001$), and negatively associated with algorithmic cocoon tightness ($r = -0.45$, $p < 0.001$). Algorithmic cocoon tightness is strongly negatively associated with perceived agenda diversity ($r = -0.65$, $p < 0.001$). All correlations are consistent with the proposed model structure.

Table 5 Pearson Correlation Matrix (N=200)

Variable	1 Structural Conditions	2 Digital Capacities	3 Behavioral Practices	4 Cocoon Tightness	5 Agenda Diversity
1 Structural Conditions	1				
2 Digital Capacities	0.52**	1			
3 Behavioral Practices	0.48**	0.60**	1		
4 Cocoon Tightness	−0.45**	−0.48**	−0.56**	1	
5 Agenda Diversity	0.42**	0.50**	0.45**	−0.65**	1

Note: ** $p < 0.001$

Correlations supported the hypothesized chain: structural conditions → digital capacities → behavioral practices → cocoon tightness → agenda diversity, all highly significant.

Significant group differences are observed across key variables.

Urban–Rural Differences (Table 6)

Urban users score higher in digital capacities ($t = 4.83$, $p < 0.001$) and behavioral practices ($t = 4.21$, $p < 0.001$), and report lower cocoon tightness ($t = -5.02$, $p < 0.001$).

Table 6 Urban–Rural Differences (N=200)

Variable	Urban (M±SD)	Rural (M±SD)	t	p
Digital Capacities	3.92±0.58	3.15±0.59	4.83	<0.001
Behavioral Practices	3.80±0.63	3.05±0.61	4.21	<0.001
Cocoon Tightness	3.02±0.62	3.81±0.53	−5.02	<0.001
Agenda Diversity	3.55±0.67	2.78±0.70	4.56	<0.001

Educational Differences (Table 7)

Higher education is associated with higher digital capacities ($F = 8.74$, $p < 0.001$) and lower cocoon tightness ($F = 6.85$, $p < 0.001$).

Table 7 Educational Differences (N=200)

Variable	≤High School	Vocational	≥Bachelor's	F	p
Digital Capacities	2.51±0.55	3.18±0.52	3.85±0.48	8.74	<0.001
Behavioral Practices	2.38±0.58	3.02±0.59	3.71±0.57	7.91	<0.001
Cocoon Tightness	3.95±0.49	3.48±0.56	2.99±0.53	6.85	<0.001

Occupational Differences (Table 8)

Significant differences are observed across occupational groups ($p < 0.001$), with professional groups showing lower cocoon tightness than manual labor groups.

Table 8 Occupational Differences (N=200)

Variable	Student	Civil Servant	Corporate	Self-employed	Manual	F	p
Digital Capacities	3.20	3.52	3.46	2.79	2.54	5.93	<0.001
Behavioral Practices	2.91	3.19	3.13	2.52	2.37	6.12	<0.001
Cocoon Tightness	3.25	2.92	3.00	3.65	3.83	6.12	<0.001

Linguistic Differences (Table 9)

Russian-speaking users report higher behavioral practices and agenda diversity than Kazakh-speaking users ($t = 2.88–3.76$, $p < 0.01$).

Table 9 Linguistic Differences (N=200)

Variable	Russian (M±SD)	Kazakh (M±SD)	t	p
Agenda Diversity	3.51±0.65	2.90±0.73	3.76	<0.001
Behavioral Practices	3.70±0.62	3.22±0.71	2.88	<0.01

Table 10 presents hierarchical regression results (DV: algorithmic cocoon tightness). After controlling for demographic variables, digital capacities ($\beta = -0.417$, $p < 0.001$) and behavioral practices ($\beta = -0.463$, $p < 0.001$) show significant negative effects.

The full model explains 51.2% of variance (Adjusted $R^2 = 0.512$), with

incremental explanatory power across steps (ΔR^2 = 0.082–0.191–0.143–0.096, all $p < 0.001$).

Table 10 Hierarchical Regression (DV: Algorithmic Cocoon Tightness, N=200)

Variable	Model 1	Model 2	Model 3	Model 4
Age	0.072	0.058	0.045	0.038
Gender	−0.065	−0.051	−0.043	−0.035
Daily Online Time	0.051*	0.042	0.033	0.029
Education	—	−0.287***	−0.173**	−0.112*
Residence	—	−0.253* *	−0.158**	−0.105*
Digital Capacities	—	—	−0.417***	−0.217***
Behavioral Practices	—	—	—	−0.463***
F	18.72***	25.34***	28.65***	30.16***
Adjusted R2	0.082	0.273	0.416	0.512
ΔR2	0.082*	0.191***	0.143***	0.096***

Note: ∗p<0.05, ∗∗p<0.01, ∗∗∗p<0.001

PROCESS Model 6 results (Table 11) confirm all indirect effects as significant based on 5,000 bootstrap samples.

The full serial mediation pathway is supported:

Structural Conditions → Digital Capacities → Behavioral Practices → Algorithmic Cocoon Tightness → Perceived Agenda Diversity.

All 95% bootstrap confidence intervals exclude zero, confirming statistical significance.

Table 11 Serial Mediation Results (N=200)

Path	Indirect Effect	SE	95% CI	Significance
Structural → Digital → Agenda	0.083	0.031	[0.070, 0.189]	Significant
Structural → Digital → Behavioral → Agenda	0.117	0.027	[0.048, 0.153]	Significant
Structural → Digital → Behavioral → Cocoon → Agenda	0.164	0.023	[0.039, 0.131]	Significant

Table 12 reports robustness results. The high-vulnerability group shows significantly higher cocoon tightness ($t = 4.97$, $p < 0.001$) and lower agenda diversity ($t = 5.04$, $p < 0.001$), while the control group demonstrates higher proactive behavioral adjustment.

Table 12 Shadow Comparative Experiment Results (N=123)

Variable	High-Risk (n=62)	Control (n=61)	t	p
Algorithmic Homogeneity	4.02±0.55	3.11±0.60	4.97	<0.001
Proactive Correction	2.28±0.59	3.47±0.64	5.31	<0.001

Variable	High-Risk (n=62)	Control (n=61)	t	p
Cocoon Tightness	3.91±0.54	2.98±0.61	−5.76	<0.001
Agenda Diversity	2.71±0.72	3.58±0.66	5.04	<0.001

Table 13 summarizes hypothesis testing results. All hypotheses (H1–H6) are supported.

Table 13 Hypothesis Testing

Hypothesis	Statement	Result
H1	Structural → Digital (+)	Supported
H2	Digital → Behavioral (+)	Supported
H3	Behavioral → Cocoon (−)	Supported
H4	Digital → Cocoon (−)	Supported
H5	Cocoon → Agenda (−)	Supported
H6	Serial mediation path	Supported

DISCUSSION AND THEORETICAL ANALYSIS

This study demonstrates that algorithmic cocoons are not purely technological outcomes but structurally embedded products shaped by social inequality, digital capacities, and user behaviors. Although algorithmic recommendation systems are designed for personalization, they are not value-neutral; rather, they extend pre-existing social inequalities into digital environments by producing differentiated patterns of information exposure. Structural disparities in education, urban–rural residence, occupation, and socioeconomic status generate systematic differences in digital capacities and information access. Groups with higher education and stronger digital skills are more capable of recognizing algorithmic mechanisms, adjusting preferences, and diversifying information sources, thereby maintaining more open informational environments. In contrast, disadvantaged groups are more dependent on default algorithmic feeds and are more likely to remain within homogeneous information flows. Behavioral practices further function as a key transmission mechanism, as digital capacities only translate into open informational environments through proactive searching, cross-platform engagement, and preference adjustment. Accordingly, algorithmic cocoon formation follows a sequential pathway-structural conditions → digital capacities → behavioral practices → informational outcomes-

through which social inequality is reproduced in digital form.

The findings confirm that the digital divide has evolved from a simple access gap into a multi-dimensional structure encompassing capacities, behaviors, and informational outcomes. A clear capacity gap exists, with urban and highly educated users demonstrating stronger abilities in information searching, verification, and algorithmic awareness, while rural and less educated users rely more heavily on default systems. These capacity differences translate into behavioral stratification, where digitally skilled users engage more actively in information exploration, whereas less skilled users exhibit more passive consumption patterns. These disparities ultimately culminate in informational outcome gaps, as even within the same platform environment, different groups experience significantly different levels of agenda diversity. The digital divide therefore becomes an informational divide that determines who is exposed to broader societal issues and who remains constrained within narrow information flows. Rather than eliminating inequality, algorithms amplify it by converting offline structural disparities into online informational stratification.

Extending agenda-setting theory to algorithmic environments, this study shows that agenda formation is no longer centralized but decentralized and structurally stratified. Algorithmic systems personalize content delivery and fragment shared public agendas into individualized information streams, resulting in agenda diversity being unevenly distributed according to digital capacities and behavioral engagement. In Kazakhstan, this stratification is further intensified by bilingualism: Russian-speaking users tend to access broader transnational content, while Kazakh-speaking users are more concentrated within localized informational environments. Consequently, parallel and partially disconnected informational agendas emerge, indicating that language functions as a structural variable in algorithmic agenda formation.

These findings suggest several implications. Platform governance should move beyond efficiency-driven personalization toward balancing relevance and informational diversity, incorporating mechanisms that reduce algorithmic

homogeneity. Digital literacy should be strengthened, particularly in algorithmic awareness, verification skills, and cross-platform information practices. Policy attention should shift from digital access to informational equity, ensuring equal opportunities for diverse information exposure. In addition, multilingual environments require improved cross-linguistic infrastructure, including translation systems and interlingual recommendation mechanisms. However, this study is subject to limitations, including a relatively small sample size that restricts generalizability, the cross-sectional design that limits causal inference, reliance on self-reported data that may introduce bias, and an individual-level focus that overlooks platform political economy structures. Future research should address these limitations through larger cross-national samples, longitudinal or behavioral trace data, objective diversity measures, and platform-level analyses. Overall, this study integrates digital divide and agenda-setting theories to demonstrate that algorithmic cocoons are structurally embedded phenomena shaped by social inequality and mediated through digital capacities and behavioral practices. The digital divide has thus evolved into a layered structure of capacity, behavioral, and informational inequalities, while algorithmic systems simultaneously reproduce and amplify these disparities, shaping differentiated informational environments across social groups.

References

1. Pariser E. The Filter Bubble: What the Internet Is Hiding from You. New York: Penguin Press, 2011.

2. Sunstein C. R. Republic.com 2.0. Princeton: Princeton University Press, 2009.

3. van Dijk J. The Deepening Divide: Inequality in the Information Society. London: Sage Publications, 2005.

4. van Dijk J. Digital Inequality. Cambridge: Polity Press, 2020.

5. McCombs M. E., Shaw D. L. The agenda-setting function of mass media // Public Opinion Quarterly. 1972. Vol. 36, No. 2. P. 176–187. DOI: 10.1080/00335637209383179.

6. Hargittai E. Digital literacy and the digital divide // Journal of Communication. 2018. Vol. 68, No. 1. P. 1–22.