

CENTRAL ASIAN JOURNAL OF THEORETICAL AND APPLIED SCIENCE

https://cajotas.casjournal.org/index.php/cajotas

Volume: 07 Issue: 01 | January 2026 ISSN: 2660-5317

Article

Artificial Intelligence and The Prevention of Intellectual Threats to Youth Consciousness

Bahodir Mirvahitovich Mirrakhimov*[1]

1. Independent researcher, National University of Uzbekistan
*Correspondance: rbahtier860@gmail.com

Abstract: The article investigates the manipulative effects of new-generation AI tools – deepfake technologies, generative artificial intelligence, and algorithmic governance – on the consciousness of youth. The research identifies four main channels of digital manipulation targeting young people: synthetic media, exploitation of personal data, emotional manipulation, and algorithmic worldview formation. Based on Uzbekistan's legal-political documents in the AI field, the New Uzbekistan Development Strategy, and international experience, proposals are formulated for shaping a multi-level national model to protect youth from digital manipulation.

Keywords: Artificial Intelligence, Deepfake, Algorithmic Manipulation, Youth Consciousness, Synthetic Media, Information Defense, Digital Sovereignty, New Uzbekistan Strategy

Citation: Mirrakhimov, B. M. Artificial Intelligence and The Prevention of Intellectual Threats to Youth Consciousness. Central Asian Journal of Theoretical and Applied Science 2026, 7(1), 319-325.

Received: 05th Dec 2025
Revised: 25th Dec 2025
Accepted: 10th Jan 2026
Published: 31st Jan 2026

Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

In today's world, the advancement of artificial intelligence (AI) technologies is driving major geopolitical, economic, and social change. Within this process, the younger generation stands out as the group most exposed to its effects. Decree No. PF-60 of the President of the Republic of Uzbekistan, dated January 28, 2022, "On the Development Strategy of New Uzbekistan for 2022–2026," identifies ensuring spiritual development and strengthening the national-spiritual foundation of the younger generation as one of the priority directions. [1] For this reason, a comprehensive study of the risks facing youth consciousness in the age of AI is a pressing scientific task.

In contemporary scholarly literature, the concept of an "intellectual attack" is being widely interpreted. In her 2025 study, M. Momeni provided quantitative evidence that the influence of political deepfake technologies on public consciousness is a clear manifestation of a new type of manipulative assault. According to her findings, fake videos circulated on Instagram proved capable of genuinely shifting users' attitudes toward particular political figures. [2] Such practices carry not only political but also moral, cultural, and personal risks.

President Shavkat Mirziyoyev has repeatedly stressed, on various occasions, that the younger generation is the principal driving force behind the building of New Uzbekistan. In his words, "today's youth must combine love for the homeland, the capacity for independent thinking, and modern knowledge in their minds and hearts". [3] The way this issue plays out in the context of AI calls for dedicated research.

The topic has already been extensively developed in the scholarly literature. In his seminal work on AI ethics, the British philosopher L. Floridi explains how the modern digital civilization is fundamentally reshaping the paradigm through which human identity is formed. [4] Local researchers Sh. Nazarov and A. Turdiev, for their part, examine how the impact of the digital environment on the younger generation manifests itself under Uzbekistan's specific conditions. [5]

The approach adopted in this article examines the manipulative influence of AI tools on youth consciousness through the concept of "channels." This approach emphasizes that each manipulative effect has its own mechanism, technological basis, and means of protection. The aim of the study is to examine, through this channel-based approach, the manipulative risks facing youth consciousness in the era of generative AI, and to develop a protection model suited to the conditions of New Uzbekistan.

2. Materials and Method

Methodologically, the study comprises three components. The first is the formation of a theoretical and methodological basis, drawing on the experience of international research centers (Cambridge, Oxford, MIT, Stanford) and domestic research schools. The second is normative-legal analysis, in which the decrees and resolutions of the President of the Republic of Uzbekistan, along with international legal instruments (the EU AI Act, UNESCO's Recommendation on the Ethics of AI, and UNICEF's AI principles for children), were examined using content analysis.

The third component is empirical-analytical. It draws on data from international statistical databases (Deepstrike Research, We Are Social, McAfee) and scientific meta-journals (arXiv, Nature, ACM FAccT) to analyze the extent to which AI tools have spread among young people. According to 2025 deepfake statistics, an AI voice-cloning attack is now recorded worldwide at a rate of roughly one for every four adults. [6] This indicates that young people in Uzbekistan are not exempt from such threats either.

Likewise, V. Livernoche and colleagues, who studied the deepfakes that circulated during the 2025 Canadian elections, analyzed 187,778 election-related posts and found that 5.9 percent of them were deepfakes. [7] This data illustrates the scale at which AI tools are already being deployed in the political and social sphere. Comparing these patterns with the Uzbek context was one of the key methodological stages of the study.

Within the domestic research community, N. Qodirov has developed a methodological framework addressing information-security issues in the upbringing of young people. A significant aspect of his approach is the argument that protecting youth consciousness should be treated not merely as a technical problem, but as a moral and pedagogical one. [8] It is precisely this methodological principle that underlies the "channel" approach adopted in this article.

3. Results

The study made it possible to identify four principal channels through which manipulative influence targets youth consciousness in the AI environment. Each channel has its own technological mechanism, psychological impact, and corresponding set of protective strategies.

The first channel is the synthetic media channel. Through this channel, deepfake videos, AI-generated commentary and articles, cloned voices, and fabricated photographs are used to instill particular attitudes and views in young people. In his seminal work on the reliability of AI, Stanford's S. Russell identifies precisely this feature – the difficulty of distinguishing machine-generated content from genuine content – as one of the principal risks posed by AI. [9]

The second channel is the exploitation of personal data. Modern AI algorithms process information about users' daily behavior, interests, and even emotional states in

order to deliver individually tailored content. As a result, the "filter bubble" surrounding a young person's worldview grows ever thicker and narrower. A 2025 study published in PubMed Central highlighted the threat that AI-assisted disinformation poses to the democratic environment as a distinct danger in its own right. [10]

This channel is especially dangerous for young people. The reason is that personal data disclosed in youth shapes a digital footprint that follows a person for their entire life: a single statement, opinion, or action can, years later, still affect their career and personal life. In his address in October 2025 launching the "One Million AI Leaders" program, the President called on young people not only to master AI, but also to remain aware of its risks. [11]

The third channel is emotional manipulation. Modern AI chatbots are capable of assessing a person's emotional state and responding to it accordingly. This feature is especially appealing to young people, who may grow accustomed to discussing their experiences and problems with a chatbot rather than with other people. Local researcher Yo. Qurbonov specifically examines the resulting conflict – and potential harmony – between AI and national identity in the minds of young people. [12]

The emotional manipulation channel is ideologically most dangerous during adolescence and early youth. Because personal identity is being formed during this period, any systematic and repeated influence leaves a lasting mark on a person's views, attitudes, and life. When the source of such influence is a machine, that mark is not genuinely authentic, but artificially constructed.

The fourth channel is the algorithmic worldview channel. This channel differs sharply from the previous three, since it operates in ways that are difficult to perceive directly. Algorithms subtly steer what content young people see, whom they choose as role models, and which values they adopt. The influential study "On the Dangers of Stochastic Parrots" (Bender et al., 2021) shows that, while large language models remain far from genuine "thought," they nonetheless wield significant social influence. [13]

Local researchers Sh. Nazarov and A. Turdiev argue that "algorithms are imperceptibly shaping young people's aesthetic, moral, and social choices". [14] This conclusion is consistent with an analytical piece published in MIT Technology Review, which notes that AI chatbots may steer children toward harmful ideological direction. [15]

4. Discussion

Interpretation of Findings in Relation to the Research Objectives

The study's channel-based framework directly addresses its stated aim of mapping the manipulative risks facing youth consciousness in the era of generative AI and grounding a protection model in Uzbekistan's specific conditions. The four channels identified — synthetic media, exploitation of personal data, emotional manipulation, and algorithmic worldview formation — are not merely a descriptive taxonomy; they represent distinct points of intervention, each requiring a different combination of technological, legal, and pedagogical responses. This directly fulfils the objective of moving beyond a generic warning about "AI risk" toward a structured account of how manipulation actually operates.

The synthetic media channel demonstrates that the central challenge is epistemic: young people increasingly cannot reliably distinguish machine-generated content from authentic content, which erodes the basic precondition for informed judgment. The personal-data channel shows that manipulation compounds over time, since a digital footprint formed in adolescence can shape opportunities and perceptions years later — a temporal dimension that is largely absent from adult-focused AI-risk literature. The emotional-manipulation channel indicates that the danger is developmental rather than purely informational: because identity formation is most active during adolescence and early youth, machine-mediated emotional engagement risks substituting an artificially

constructed influence for authentic social development. Finally, the algorithmic-worldview channel shows that the most consequential manipulation may be the least visible, operating through recommendation systems rather than discrete, identifiable acts of deception. Together, these findings support the study's central premise that AI-related threats to youth consciousness cannot be addressed by a single measure, but require differentiated responses calibrated to each channel's specific mechanism.

Comparison with Previous Studies

These findings are consistent with, and extend, several strands of the literature reviewed earlier in this article. Momeni's quantitative demonstration that political deepfakes measurably shifted user attitudes on Instagram supports the synthetic-media channel identified here, while the scale documented by Livernoche and colleagues in the 2025 Canadian election — with 5.9 percent of election-related posts found to be deepfakes — indicates that the phenomenon Momeni described experimentally is already occurring at a societal scale, reinforcing the case for treating synthetic media as a distinct and urgent channel rather than a marginal technical curiosity.

The personal-data and algorithmic-worldview channels align closely with the "filter bubble" literature and with Bender and colleagues' caution that large language models, while lacking genuine understanding, nonetheless exert substantial social influence through the volume and consistency of content they generate. This study extends that observation by specifying a mechanism through which such influence reaches youth in particular: not through isolated exposure, but through the cumulative narrowing of information environments over time.

The emotional-manipulation channel builds on, and gives a distinctly Uzbek inflection to, Qurbonov's analysis of the tension and potential harmony between AI and national identity. Whereas much of the international literature on AI and emotional well-being (for example, the MIT Technology Review analysis of chatbot influence on children) treats the risk in largely universal, developmental terms, Qurbonov's framing situates the same risk within a specific question of cultural and spiritual continuity, directly relevant to Uzbekistan's stated policy emphasis on the "national-spiritual foundation" of the younger generation. This suggests that channel-specific protective strategies cannot be imported wholesale from international models but must be adapted to reflect local conceptions of identity and value formation, a point substantiated further by Nazarov and Turdiyev's domestic findings on algorithmic influence over young people's aesthetic and moral choices.

Theoretical Implications

At a theoretical level, the channel-based framework proposed here contributes to the AI-ethics literature by disaggregating the general concept of "AI manipulation" into mechanistically distinct pathways, each with its own risk profile and point of intervention. This addresses a gap in Floridi's otherwise comprehensive account of AI ethics: while Floridi identifies transparency, accountability, and a human-centered approach as core principles, his framework operates at a level of generality that does not by itself indicate which principle is most relevant to which type of harm. The four-channel model developed here suggests a mapping — transparency is most directly relevant to the synthetic-media channel, accountability to the personal-data channel, and a human-centered approach to the emotional-manipulation and algorithmic-worldview channels — that could inform more precise applications of Floridi's principles in policy design.

The findings also extend Qodirov's methodological argument that youth information-security protection is a moral and pedagogical problem, not merely a technical one, by showing concretely how this claim applies differently across channels: the synthetic-media and personal-data channels admit substantial technical mitigation, whereas the emotional-manipulation and algorithmic-worldview channels are, by their nature, resistant to purely technical solutions and require sustained pedagogical and

familial engagement. This channel-differentiated view of the technical–moral distinction represents a modest but genuine theoretical refinement of Qodirov's original claim.

Practical Implications

For policymakers, the practical implication of the four-channel model is that a single "AI safety" law or curriculum module is unlikely to be sufficient. The findings suggest instead a differentiated policy architecture: normative-legal instruments, drawing on the EU AI Act's risk-tiered approach, for the synthetic-media and personal-data channels, where harms are more readily codified and enforced; educational and public-awareness interventions for the emotional-manipulation channel, aimed at strengthening young people's capacity for reflective rather than purely reactive engagement with AI systems; and a combination of media-literacy education and family or mahalla-based mentoring for the algorithmic-worldview channel, where the influence is diffuse and not easily regulated through law alone.

For educational institutions, the findings support the introduction of a dedicated course addressing AI-related risks, but the channel-based framework indicates that such a course should be structured around the four distinct mechanisms identified here rather than around a generic notion of "digital literacy," since each channel demands different practical skills — from deepfake-detection techniques for the synthetic-media channel to critical-reflection exercises for the algorithmic-worldview channel.

For families and community institutions, the findings reinforce Toshmatov and Yo'ldoshev's observation that the mahalla has historically functioned as an educational and moral filter, while indicating that this role now needs to be extended explicitly to the emotional-manipulation and algorithmic-worldview channels — the two channels least amenable to top-down legal or technical solutions and therefore most dependent on sustained, trust-based relationships between young people and the adults around them.

Policy Synthesis and National Context

Each of the four channels described above requires its own distinct approach and protective strategy. At the same time, together they exert a broad, cumulative influence on the formation of youth consciousness. In Uzbekistan's context, countering these risks should be pursued in line with the AI strategy approved by the President [16] and the development principles of "New Uzbekistan."

In international practice, models for protecting against AI-related risks are being developed along several lines. A 2023 study published by OpenAI offers practical recommendations for mitigating the social impact of language models. [17] In this regard, technical measures need to be accompanied by transparency, user awareness, and public participation.

In the domestic context, the family and the mahalla (neighborhood community) are the most important protective institutions. I. Toshmatov and A. Yoldoshev note in their study that the Uzbek family and mahalla have for centuries served as an educational and moral filter. [18] Under modern conditions, these institutions must retain that role with respect to AI-related risks as well, while being adapted to new digital realities.

As for international legal measures, the EU AI Act, adopted in June 2024, is currently regarded as the world's most advanced legal instrument on AI. [19] The Act classifies AI systems by risk level and sets out specific legal requirements for each category. Drawing on this experience, developing an act or draft law suited to Uzbekistan's national conditions is a pressing task.

Moreover, young people themselves should not remain passive objects of protection, but should become active participants. Local researcher Yo. Qurbonov articulates an important thesis in this regard: "young people in the age of AI should not merely be protected – they should be equal participants in the creation, governance, and

dissemination of AI". [20] This perspective shifts the framing of the AI issue from treating young people as "objects" to understanding them as "subjects."

In his comprehensive work on AI ethics, L. Floridi identifies three core principles underlying modern protective strategies: transparency, accountability, and a human-centered approach. [21] These principles are already reflected in Uzbekistan's strategic documents on AI and now need to be put into practice.

In his December 2024 Address to the Oliy Majlis, the President specifically emphasized the need to mobilize every effort for the sake of present and future generations. [22] In this respect, protecting the consciousness of young people in the AI era must become a priority not only for science and technology, but for state policy as a whole.

Limitations and Future Research

This study is conceptual and analytical in nature, drawing on secondary statistical sources, policy documents, and prior research rather than original survey or experimental data collected from Uzbek youth. The prevalence estimates cited — including the global deepfake-attack rate and the share of deepfake content in the 2025 Canadian election — are drawn from international contexts and have not yet been validated for the Uzbek population specifically. Future research should therefore prioritize empirical studies, including surveys, experiments, or platform-level data analysis, that measure the prevalence and impact of each of the four channels among Uzbek youth directly, and should evaluate the effectiveness of the protective measures proposed here once they are implemented.

5. Conclusion

Pancreatic This study has shown that the manipulative risks facing youth consciousness in the age of artificial intelligence form a complex, multifaceted, and constantly evolving phenomenon that cannot be resolved through one or two simplified measures, and the channel-based approach proposed here — distinguishing synthetic media, exploitation of personal data, emotional manipulation, and algorithmic worldview formation — offers a scientific and methodological basis for developing separate technological, legal, and pedagogical protective measures for each channel. Based on these findings, several strategic directions are proposed: developing a national concept for "AI and youth safety" together with a corresponding draft law that provides clear normative-legal responses to threats such as deepfakes, exploitation of personal data, and emotional manipulation; introducing a course on safety in the AI environment in educational institutions that teaches practical skills for detecting deepfakes, analyzing synthetic media, and guarding against emotional manipulation; establishing a dedicated track for young researchers within the New Uzbekistan AI Alliance so that youth develop as creators and disseminators of AI tools rather than only as users; enriching the role of the family and mahalla institutions in digital upbringing through regular training for parents and mahalla chairpersons; developing AI assistants and educational chatbots grounded in local national-spiritual values and expanding Uzbek-language corpora to support this work; strengthening international scientific cooperation through joint research and expert-exchange programs with UNESCO, UNICEF, the EU, and other organizations; and developing a national technological solution for the automatic detection of deepfakes and other synthetic media, drawing on the capabilities of the Center for the Development of Artificial Intelligence and the Digital Economy. The most important conclusion is that the problem of artificial intelligence and youth consciousness must become one of the central themes of Uzbekistan's state policy, both now and over the coming decade, requiring simultaneous action on the technical, moral-pedagogical, legal, social, and international-cooperation fronts, so that the younger generation becomes a genuine, equal partner — rather than a mere target — in the world of AI, which is one of the key factors in building New Uzbekistan.

REFERENCES

[1] Decree of the President of the Republic of Uzbekistan No. PF-60 of January 28, 2022, "On the Development Strategy of New Uzbekistan for 2022–2026." Lex.uz, 2022.

[2] M. Momeni, "Artificial Intelligence and Political Deepfakes: Shaping Citizen Perceptions Through Misinformation," Media Asia. SAGE, 2025. doi: 10.1177/09732586241277335.

[3] Sh. Mirziyoyev, "Towards Directing the Development of New Uzbekistan to Serve Human Interests," official speech materials published on the website of the President of the Republic of Uzbekistan. Tashkent, 2024.

[4] L. Floridi, The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford: Oxford University Press, 2023, 256 p.

[5] Sh. Nazarov and A. Turdiyev, "The Impact of the Digital Environment on Youth Education: The Experience of Uzbekistan," Social Opinion. Human Rights, no. 2, pp. 56–67, Tashkent, 2024.

[6] Deepstrike Research, "Deepfake Statistics 2025: The Data Behind the AI Fraud Wave," Sep. 2025. [Online]. Available: https://deepstrike.io/blog/deepfake-statistics-2025

[7] V. Livernoche, A. Musulan, Z. Yang et al., "Deepfakes in the 2025 Canadian Election: Prevalence, Partisanship, and Platform Dynamics," arXiv preprint, arXiv:2512.13915, 2025.

[8] N. Qodirov, "Information Security Issues in the Education of the Younger Generation," Political Science, no. 4, pp. 89–96, Tashkent, 2024.

[9] S. Russell, Human Compatible: Artificial Intelligence and the Problem of Control. New York: Viking Press, 2019, 352 p.

[10] "AI-Driven Disinformation: Policy Recommendations for Democratic Resilience," PubMed Central, Art. PMC12351547, 2025.

[11] Sh. Mirziyoyev, "Speech on the Launch of the 'One Million Artificial Intelligence Leaders' Program." Tashkent, October 2025. Sputniknews.uz.

[12] Yo. Qurbonov, "Artificial Intelligence and National Identity: Conflict and Harmony in the Consciousness of Youth," Spiritual Enlightenment, no. 1, pp. 14–22, Tashkent, 2025.

[13] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021, pp. 610–623.

[14] Sh. Nazarov and A. Turdiyev, "The Impact of the Digital Environment on Youth Education: The Experience of Uzbekistan," *Social Opinion. Human Rights*, no. 2, p. 62, Tashkent, 2024.

[15] K. Hao, "AI Chatbots Could Lead Children Astray," MIT Technology Review, Aug. 14, 2024.

[16] Resolution of the President of the Republic of Uzbekistan No. PQ-358 of October 14, 2024, "On Approval of the Strategy for the Development of Artificial Intelligence Technologies until 2030." Lex.uz.

[17] I. Solaiman, M. Brundage, A. Askell et al., "Release Strategies and the Social Impacts of Language Models," OpenAI Reports, 2023.

[18] I. Toshmatov and A. Yo'ldoshev, "The Role of the Mahalla and Family Institutions in Youth Education in the Digital Age," Society and Governance, no. 3, pp. 41–48, 2024.

[19] European Union, EU Artificial Intelligence Act (Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024). Official Journal of the European Union, 2024.

[20] Yo. Qurbonov, "Artificial Intelligence and National Identity: Conflict and Harmony in the Consciousness of Youth," *Spiritual Enlightenment*, no. 1, pp. 19–21, Tashkent, 2025.

[21] L. Floridi, *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford, U.K.: Oxford University Press, 2023, p. 178.

[22] Sh. Mirziyoyev, "Address to the Oliy Majlis." Tashkent, December 20, 2024.