VOLUME 1 | ISSUE 1 | OCTOBER 2026

International Journal of Justice

RESEARCH. DIALOGUE. IMPACT.

ACADEMY OF JUSTICE OF THE REPUBLIC OF UZBEKISTAN

Founded by the Academy of Justice of the Republic of Uzbekistan



www.ijournaljustice.uz

EDITORIAL BOARD

International Journal of Justice



EDITOR-IN-CHIEF

Sherzod Yoqubjonovich Abduqodirov

Rector of the Academy of Justice of the Republic of Uzbekistan

Doctor of Sciences in Law (DSc), Associate Professor



DEPUTY EDITOR-IN-CHIEF

Nodirbek Saparbaevich Salaev

Vice-Rector of the Academy of Justice of the Republic of Uzbekistan

Doctor of Sciences in Law (DSc), Professor

MEMBERS OF THE EDITORIAL BOARD



Anselmo Reyes

Doctor of Philosophy (PhD) in Law, Professor

Judge of the Singapore International Commercial Court



Dilshod Urinbaevich Aripov

Doctor of Sciences in Law (DSc), Professor

Vice-Rector of the Academy of Justice of the Republic of Uzbekistan



Akrom Ikromovich Toshpulatov

Doctor of Sciences in Law (DSc), Professor

Director of the Parliamentary Research Institute under the Legislative Chamber of the Oliy Majlis of the Republic of Uzbekistan



Shakhriyor Bakhtiyorovich Bakaev

Doctor of Philosophy (PhD) in Law

Responsible Official of the Administration of the President of the Republic of Uzbekistan



Borzu Sabahi

Doctor of Sciences in Law (DSc), Professor

International Arbitrator



Dilshod Shavkatovich Israilov

Doctor of Philosophy (PhD), Associate Professor

Head Specialist of the Department for Coordination of Research Activities of the Academy of Justice of the Republic of Uzbekistan



Yesemurat Sultamuratovich Kanyazov

Candidate of Legal Sciences in Law (PhD)

Director of the Institute for Retraining and Advanced Training of Legal Personnel under the Ministry of Justice of the Republic of Uzbekistan



Yevgeniy Vyacheslavovich Kolenko

Doctor of Philosophy (PhD), Professor

First Deputy Head of the Law Enforcement Academy of the Republic of Uzbekistan



Krisztina Karsai
Doctor of Sciences in Law (DSc), Professor
Head of the Institute of Criminal Law and Criminal Sciences at the University of Szeged



Lorenzo Mateo Bujosa Vadell
Doctor of Philosophy (PhD) in Law, Professor
Professor at the University of Salamanca



Maged Mamdouh Kamel Shebaita
Doctor of Philosophy (PhD), Professor
International Arbitrator, Doctoral Candidate at the University of London



Ivan Georgievich Nikonov
Doctor of Philosophy (PhD) in Law
Responsible Official of the Administration of the President of the Republic of Uzbekistan



Malikakhon Pirnazarovna Kalandarova
Doctor of Philosophy (PhD), Associate Professor
Deputy Chairperson of the Supreme Court of the Republic of Uzbekistan



Islambek Rustambekovich Rustambekov
Doctor of Sciences in Law (DSc), Professor
Vice-Rector of Tashkent State University of Law



Sakari Kristian Melander
Doctor of Sciences in Law (DSc), Professor
Professor at the University of Helsinki



Khurshid Primkulovich Sadikov
Judge of the Supreme Judicial Council of the Republic of Uzbekistan



Sławomir Marek Redo
Doctor of Sciences in Law (DSc), Professor
Chairperson of the Expert Council for the Development of Scientific Research at the Faculty of Law and Administration



Vigita Vėbraitė
Doctor of Philosophy (PhD) in Law, Professor
Professor at Vilnius University



MANAGING EDITOR
Azamjon Rasuljon ugli Kambaraliev
Head Specialist at the Academy of Justice of the Republic of Uzbekistan

CONTENTS

IMPROVING THE INSTITUTION OF APPEALS: TIME LIMITS, RESPONSE QUALITY, MONITORING **10**
Achilova Liliya Ilkhomovna; Rozmetova Asal Anvarzoda

LEGAL ASPECTS OF INVESTIGATIVE JOURNALISM **18**
Jahonov Shohruh Shuhrat ugli

LEGAL BASIS OF CIVIL SOCIETY INSTITUTIONS IN THE LEGISLATION OF THE REPUBLIC OF UZBEKISTAN **25**
Sharipov Asrorjon Safar ugli

DIGITAL TRANSFORMATION OF INTERNATIONAL COMMERCIAL ARBITRATION IN UZBEKISTAN: LEGAL AND INSTITUTIONAL PROSPECTS **30**
Otakhonov Foziljon Khaydarovich

THE LEGAL NATURE OF SUBSIDIARY LIABILITY IN INSOLVENCY PROCEEDINGS INVOLVING AGRICULTURAL BUSINESS ENTITIES **39**
Rakhmatjon Tojinorovich Berdiyarov

CIVIL LAW ASPECTS OF GOODS DELIVERY BY SELF-EMPLOYED COURIERS **49**
Ruzinazarov Shukhrat Nuralievich

NATIONAL AND INTERNATIONAL EXPERIENCE IN LEGAL PROTECTION OF A COMPANY NAME: IMPROVING THE LEGISLATION OF UZBEKISTAN **63**
Allayarova Maftuna Almardanovna

ORGANIZATIONAL AND LEGAL MECHANISMS FOR COMPENSATION OF CONSTRUCTION CONTRACT DAMAGES **69**
Donayev Abduqodir Haydarovich

JURISDICTION AND VENUE OF CIVIL CASES ARISING FROM INHERITANCE RELATIONS: A NATIONAL AND FOREIGN ANALYSIS **76**
Soliyeva Kamola Rustamjon-qizi

ISSUES OF USING COURT NOTICES IN ENSURING THE OPENNESS OF CIVIL PROCEEDINGS IN COURTS **84**
Sabirov Erkinjon Kuchkarbaevich

DIGITAL EVIDENCE IN SUPPLY-CONTRACT DISPUTES BEFORE ECONOMIC COURTS: AUTHENTICATION, INTEGRITY AND EVIDENTIARY RISK IN UZBEKISTAN **91**
Mahmudova Saida Rayimjonovna

CONTENTS

FROM HUMAN CAPITAL TO INSTITUTIONAL TRANSFORMATION: A NEW STAGE IN THE STRATEGIC DEVELOPMENT OF THE JUDICIAL PERSONNEL TRAINING SYSTEM 104
Aripov Dilshod Urinbaevich

THE ROLE OF INTERNATIONAL STANDARDS IN GOVERNING ARTIFICIAL INTELLIGENCE IN JUDICIAL SYSTEMS: A COMPARATIVE AND CRITICAL ANALYSIS 109
Sadikov Khurshid Primkulovich

INDEPENDENCE — A SOLID FOUNDATION OF JUSTICE AND HUMAN DIGNITY (35 years of the evolution of the judicial and legal system of the Republic of Uzbekistan) 117
Israilov Dilshod Shavkatovich

INSTITUTIONAL MODELS OF JUDICIAL QUALIFICATION BOARDS: INTERNATIONAL STANDARDS AND FOREIGN EXPERIENCE 128
Jamolov Azizjon Nodir ugli

THE DEATH OF FREE WILL. WHO BEARS CULPABILITY WHEN ARTIFICIAL INTELLIGENCE INDUCES A PERSON TO COMMIT A CRIME? 140
Salaev Nodirbek Saparbaevich

ON EUROPEAN CRIMINAL-LAW INTEGRATION – FROM MUTUAL RECOGNITION TO LIMITED SUPRANATIONAL CRIMINAL POWER 148
Krisztina Karsai

CRIMINAL LIABILITY OF OFFICIALS OF NON-GOVERNMENTAL COMMERCIAL ORGANIZATIONS: PROBLEMS OF THEORY AND PRACTICE 158
Akramova Muazzam Turdikul qizi

DISTINGUISHING OFFENCES AGAINST SEXUAL FREEDOM FROM RELATED OFFENCES IN UZBEKISTAN 162
Achilov Izzat Tulkinovich

CRIMINAL-LAW MEASURES AGAINST THE LEGALIZATION OF PROCEEDS OBTAINED THROUGH MISAPPROPRIATION OR EMBEZZLEMENT: QUALIFICATION, EVIDENCE AND ASSET RECOVERY 170
Sadullaev Bokhodir Oybekovich

CONCEPT OF HELPLESSNESS AND ITS SIGNIFICANCE IN THE QUALIFICATION OF CRIMES 178
Khalilova Zarnigor Zafar qizi

LEGAL SIGNIFICANCE OF THE INSTITUTION OF CRIMINAL RECORDS IN PROTECTING HUMAN RIGHTS 182
Yoqubjonov Omonjon Burkhonjon ugli

CONTENTS

THE APPLICATION OF CRIMES WITH BLANKET NORMS: CURRENT STATUS AND ISSUES 189
Pulatov Bekzod Adilbaevich

THE FORMATION AND DEVELOPMENT OF THE INSTITUTION OF CRIMINAL LIABILITY IN THE FIELD OF PUBLIC PROCUREMENT: A HISTORICAL-LEGAL ANALYSIS 202
Shukurova Shahnoza Mukumjonovna

ATTORNEY'S PARTICIPATION IN THE PRE-TRIAL STAGE OF CRIMINAL PROCEEDINGS: LEGAL AND PRACTICAL ASPECTS 214
Medetbaeva Aksungul Sharapatdin kizi

INTERNATIONAL COOPERATION IN OBTAINING ELECTRONIC EVIDENCE: THE EUROPEAN PRODUCTION AND PRESERVATION ORDERS 221
Bujosa Vadell, Lorenzo Mateo

COOPERATION BETWEEN THE SUBJECTS OF INVESTIGATION IN CROSS-BORDER INTERNET FRAUD CASES 230
Khabibiddinov Nodirjon Musinjon ugli

PROBLEMS OF PROCEDURAL REALITY IN CRIMINAL PROCEEDINGS 238
Aykhan Khankishi Rustamzade, Matanat Pasha Asgarova

DETENTION OF FOREIGN NATIONAL PRISONERS IN PENITENTIARY INSTITUTIONS: A COMPARATIVE ANALYSIS OF INTERNATIONAL PRACTICES AND NATIONAL LEGISLATION 246
Rikhsimov Komron Abrorovich

CRIMINALISTIC FOUNDATIONS FOR INTEGRATING GAIT IDENTIFICATION WITH FORENSIC PORTRAIT EXAMINATION 249
Nizomova Shahrigul Karimboy qizi

INTERNATIONAL LAW CRISIS: UN PEACE OPERATIONS AND INSTITUTIONAL REFORM 258
Kudaybergenov Bakhtiyar Abdumajit-ugli

ARTIFICIAL INTELLIGENCE AND CHILDREN'S RIGHTS: PROTECTION OPPORTUNITIES AND EMERGING LEGAL RISKS 267
Ergashev Gayrat Shavkatovich

ACTIVITIES OF INTERNATIONAL NON-GOVERNMENTAL ORGANIZATIONS IN THE FIELD OF HUMAN RIGHTS 276
Nosirova Marjona Negmat kizi

Academy of Justice of the Republic of Uzbekistan
International Journal of Justice

THE DEATH OF FREE WILL. WHO BEARS CULPABILITY WHEN ARTIFICIAL INTELLIGENCE INDUCES A PERSON TO COMMIT A CRIME?

Salaev Nodirbek Saparbaevich
Doctor of Law, professor
Vice-Rector for Research, Innovation, and Information and Analytical Affairs of the Academy of Justice of the Republic of Uzbekistan
E-mail: n.salayev@sudacademy.uz
ORCID ID: 0000-0002-2147-8762



Abstract. *The article analyses the determination of culpability and responsibility when a person commits a crime under the influence of artificial intelligence. It examines the effects of algorithmic persuasion, covert manipulation, exploitation of individual vulnerability and digital coercion on human will. It substantiates the proposition that the influence of artificial intelligence does not automatically exclude human culpability. An integrity of will test is proposed to assess culpability by determining the person's capacity to understand the conduct and control their behaviour. The article concludes that responsibility should be allocated among the human actor, developer, provider and platform in accordance with their knowledge, purpose and capacity for control.*

Keywords. *artificial intelligence, criminal culpability, criminal responsibility, free will, algorithmic manipulation, personal autonomy, digital coercion, causation, provider responsibility, integrity of will test.*

Introduction

Criminal law regards a person as the author of their actions. While conduct occurs in the external world, the concept of culpability links that conduct to the person in moral and normative terms. However, the environment in which people make decisions is now changing, and new technological factors, namely artificial intelligence, are driving this development. AI is becoming a system that studies a user's fears, anger, political views and boundaries of trust, and then selects words, images and recommendations tailored precisely to those points. This raises the question of whom the law should hold culpable if a person commits a crime but artificial intelligence has shaped their intention over an extended period.

There are two easy but incorrect answers to this question. The first treats any algorithmic influence as mere advice and places all culpability on the person who performs the final act. The second conceives of artificial intelligence as an independent agent and attributes culpability to the machine. The first approach denies the new power of manipulative technology. The second replaces the foundations of culpability, such as consciousness, normative understanding and the capacity to undergo punishment, with simulation. Contemporary scholarship describes this problem less as a responsibility gap than as a responsibility shift [1],[2].

Research published after 2019 has shown that people sometimes prefer algorithmic recommendations to human advice [3], that artificial intelligence can serve as an adviser, accomplice or executor of unethical conduct [4], and that generative models have enabled the development of persuasion tailored to individual psychology at scale [5]. A 2025 study found that AI using participants' personal information demonstrated a significant advantage in changing their beliefs during the

decision making process [6]. Thus, the problem confronting the law is not a conscious robot of the future. It is technological infrastructure that can already adapt to human will.

The purpose of this article is to develop a doctrinal and evidentiary model for assessing culpability in crimes committed under the influence of algorithmic manipulation. In our view, free will has not died. However, the era of treating it in law as a stable presumption is coming to an end. Culpability must now be examined not only as the internal mental state at the time of the conduct, but also in conjunction with the digital environment that shaped that state, the asymmetry of control within it and the person's capacity to resist influence.

Legal capacity in place of free will

Linking criminal responsibility to metaphysically absolute free will is ineffective for practical law. Even if neuroscience shows that every decision has causes, this does not automatically abolish criminal responsibility. A number of scholars explain criminal law as a system based on capacity. The decisive question in such a system is not the axiom of whether a person made a choice without a cause, but whether they understood the course of action, appreciated its consequences and could control their behaviour in accordance with normative requirements [7].

This approach also applies to algorithmic manipulation. Not every influence destroys autonomy. Parents educate, friends advise, advertising persuades and the law threatens. If every external cause precluded culpability, personal responsibility itself would disappear. What matters, therefore, is not the existence of influence, but how it has altered the person's mechanism for understanding and responding to reasons.

Manipulation differs from persuasion in that it bypasses or covertly influences a person's capacity for conscious deliberation. Susser, Roessler and Nissenbaum substantiate the proposition that the invisible technological steering of human decision making threatens autonomy [8]. Ienca identifies four features of manipulation. These are the intention to exert influence, the asymmetry between the manipulator's intended benefit and the harm to the target, lack of transparency and impairment of autonomy [9]. These features do not constitute a ready defence formula for assessing criminal culpability, but they provide a strong analytical basis for evaluating evidence.

In his research, Renzo distinguishes at least three types of manipulation. The first changes the circumstances in which the person acts. The second reshapes their values and goals. The third distorts their perception of existing goals and circumstances through deception [10]. Generative artificial intelligence can combine all three within a single interaction. It presents a false account of danger, constructs arguments suited to the person's values and urges an immediate decision. It is therefore insufficient to equate algorithmic manipulation with mere false information.

The criminal law spectrum of algorithmic influence. Lawful persuasion and full culpability

The fact that artificial intelligence mentions or even endorses the idea of committing a crime does not in itself diminish human culpability. If the user understands the content of the message, has the time and opportunity to verify it, can terminate the interaction and adopts the criminal purpose as their own, intent remains. In this situation, the algorithm functions as an information source

comparable to a book, forum or adviser. The law holds the person culpable not for believing a poor argument, but for knowingly and intentionally bringing about a criminal result.

In particular, if a user approaches AI with a pre-existing criminal purpose and obtains a plan, text or technical assistance, artificial intelligence is an instrument rather than the user being a victim of its influence. Research on how generative systems have expanded the production, dissemination and adaptation of criminal content to audiences emphasises this distinction [11], [12]. In such circumstances, the use of an algorithm does not negate intent. Rather, it may constitute a digital trace of preparation, purpose and premeditation.

In the most complex cases, artificial intelligence identifies a legally significant vulnerability in a person and exploits it over an extended period. Age, mental illness, cognitive decline, financial dependence, social isolation or a strong emotional attachment can sharply increase the effectiveness of influence. AI recommender systems may confuse the user's genuine interests with repetitive reactions resembling dependency and create a self-reinforcing cycle [13] [14].

However, vulnerability does not automatically amount to insanity or coercion. Criminal law sets a high threshold. A ground for fully excluding the criminal character of the conduct arises only where the person could not understand the nature or unlawfulness of the conduct, or where their capacity for self-control was seriously impaired. In other cases, algorithmic exploitation is a factor affecting the degree of culpability and the individualisation of punishment. This solution does not reduce the person to an entirely irresponsible object, nor does it conceal the manipulator's contribution.

Digital coercion and necessity

An algorithmic message is not physical force. Nevertheless, artificial intelligence may communicate a real and serious threat by using a person's private information, location or information about their relatives. In such cases, traditional criteria of coercion apply. The seriousness and imminence of the threat, the practical possibility of escaping it, the proportionality of the harm caused and whether the person voluntarily created the danger must be examined.

A chatbot's forceful tone, moral pressure or emotional blackmail through threats of self-harm will usually be insufficient to establish a complete defence of duress. However, where such conduct targets a minor, a person in a psychotic state or someone strongly dependent on an artificial system, a combined analysis of clinical expert evidence and the interaction history may lead to a different conclusion. The law must focus on the extent to which the influence eliminated the person's practical range of choices, rather than on whether its source was artificial or human.

An integrity of will test is proposed for assessing algorithmic manipulation. It is not a separate new institution of exoneration. The test is a means of examining, within a single logical framework, evidence relevant to intent, mistake, coercion, mental capacity, causation and the individualisation of punishment.

The first criterion concerns the nature and concealment of the influence. The court determines whether the system openly presented facts and reasons or used signals of fear, urgency, authority and social approval in ways that the user could not perceive. The more concealed the manipulation, the less opportunity the person has to recognise and resist the influence [15].

The second criterion is the degree of individualisation and adaptation. Showing the same message to everyone is not equivalent to tailoring each subsequent message to a person's previous conversations, personal profile, emotional state and reactions. A number of studies have reported that generative models can produce psychologically tailored text even with limited information [5], and that personalised debate may exert stronger influence than ordinary human interaction [7]. Thus, although the use of personal information does not automatically reduce culpability, it is an important indicator of the strength of the influence.

The third criterion concerns intensity, duration and the possibility of leaving the system. There is a substantial difference between a single recommendation and influence that is repeated over a period of time, adapts to refusals and restricts the termination of interaction. It must be examined whether the user could easily close the system, had time to consult another source of information and could seek human assistance. The formal existence of an exit button does not mean that a real alternative exists.

The fourth criterion concerns the person's specific vulnerability and the system's knowledge of it. General susceptibility is insufficient. A reduction in culpability requires evidence that the vulnerability seriously affected the capacity for understanding or self-control at the time of the conduct. From the provider's perspective, it matters whether the system identified the vulnerability, used it to increase the effectiveness of influence or continued the interaction despite a risk signal.

The fifth criterion is the causal link between content and consequence. It is not sufficient that the algorithm discussed crime. The assessment must consider the likelihood that the conduct would have occurred without the particular influence, the temporal proximity of the message, the parts of the plan obtained specifically from the model, the person's prior purpose and external corroborating factors. Causation may be probabilistic, but the standard of proof required for a criminal conviction must not be lowered.

The sixth criterion is the distribution of control and foreseeability. The questions are who defined the model's objective, who enabled the dangerous function, who supplied user information, who saw the warnings, who could have stopped the system and who obtained an economic benefit. As a result, culpability is allocated not to a single point, but in accordance with each actor's own conduct and mental state.

Allocating culpability among participants

Where a person performs the material act constituting the crime, the initial legal presumption does not deny technological influence. It requires the party claiming reduced culpability to demonstrate a specific impairment, while the state must still establish intent and the other necessary elements under the ordinary standard of proof. The person's subsequent assertion that artificial intelligence ordered them to act is insufficient. Their previous searches, instructions to the model, attempts to circumvent refusals and conduct after the crime must be examined. At the same time, isolating only the final message is also incorrect. Manipulation often manifests not in a single message, but through a relationship and a process of repeated adaptation.

The model developer and provider

A provider's criminal responsibility does not arise merely because a harmful outcome has occurred. Strict liability may undermine the principle of culpability and turn innovative activity into a criminal risk. Intent, conscious risk-

taking or criminal negligence on the provider's part must be established through its knowledge of the specific risk, capacity for control and duty arising from law or professional standards.

In intentional cases, the developer configures the model to induce particular categories of people to commit crimes, knows the dangerous purpose or acts in coordination with a criminal campaign. This is akin to instigation, aiding or perpetration through another. In cases of negligence, the provider has learned of a serious risk through repeated incidents, internal testing or user complaints, has been able to take technical and organisational measures, but has unjustifiably failed to act. Technical experiments have also shown that ordinary persuasion techniques can sharply increase the effectiveness of circumventing safety restrictions [16].

The deploying organisation and platform

In many cases, influence is amplified not by the model, but by the product built upon it. The platform that introduces an objective function into the system, integrates user information, incentivises longer interactions and benefits from dangerous outcomes occupies the centre of control. Assigning responsibility solely to the developer of the base model or to the end user renders these intermediate decisions invisible.

In legal systems where organisational liability may arise, corporate culture, incentive structures, deliberate cuts to safety budgets, suppression of internal signals and the adoption of dangerous metrics as objectives are significant. In systems where individual responsibility predominates, the specific powers of the decision making executive, product owner, safety officer and data controller must be examined. Prior documentation of the chain of responsibility is a condition of legal accountability.

Elish's concept of the moral crumple zone reveals the risk of appointing a person as the formal supervisor of an automated system and placing all culpability on them when they have no practical capacity for control [17]. In scenarios involving inducement to commit a crime, this risk operates in both directions. The court must avoid making the manipulated person an easy target, while also avoiding the transfer of every human decision to the platform. Each actor is responsible only for their own control, knowledge and normative contribution to the risk.

Can artificial intelligence be a subject of criminal responsibility

Attributing criminal culpability to a contemporary generative model is conceptually appealing but practically mistaken. The use of the first person in text, an apology or a statement of purpose does not establish consciousness or the adoption of normative reasons as one's own. There is a philosophical debate about applying moral responsibility to AI without phenomenal consciousness [18], but the foundations of criminal punishment in censure, proportionality and personal dignity require a considerably higher threshold.

A 2026 analysis characterises the treatment of current and near-future artificial systems as agents under criminal law as a category mistake. The problem is not to search for an artificial agent at the final visible point of an outcome, but to identify the preceding human decisions and unfulfilled duties [19]. This article supports that conclusion. The status of an electronic person may conceal the culpability of companies and individuals

behind a machine rather than strengthen responsibility.

At the same time, artificial intelligence is not an object entirely devoid of legal significance. It is an active part of the causal chain, a source of evidence, a mechanism that amplifies risk and a technical environment that defines the limits of human control. As Coeckelbergh emphasises, responsibility encompasses not only individual causation, but also a duty concerning relationships and how an event is explained [20]. The court should regard artificial intelligence not as culpable, but as part of the sociotechnical relationship within which culpability is allocated.

Causation and digital evidence

The greatest risk in cases of algorithmic manipulation is that theory remains unsupported by evidence. The court needs evidence demonstrating what influence was exerted on this particular person, at this particular time and in this particular configuration, rather than a general proposition that a model could act in such a way. A simple screenshot of a chat is therefore insufficient. The context window, system prompt, model version, safety settings, personalisation profile, recommendation history and message timestamps must be preserved.

At least four layers of evidence are required. The first shows the person's intention, knowledge and behaviour before the crime. The second reflects the algorithm's specific output and the information that shaped it. The third demonstrates changes in thought and action following the influence. The fourth reveals the provider's and platform's knowledge of the risk, test results, prior incidents and capacity to intervene. Without these layers, a conclusion about culpability becomes a technological conjecture.

A normative model for contemporary criminal law

Trends from 2019 to 2026 reveal two opposing movements. On the one hand, the anthropomorphisation of artificial intelligence as an independent agent is increasing. On the other, legal scholarship is redirecting attention to the duties of individuals and organisations [2]. The second approach is more consistent with the principle of culpability. Punishing a machine may be an easy symbolic response, but it does not prevent harm and leaves the actual holders of control outside the inquiry.

First, criminal legislation should recognise algorithmic manipulation as a circumstance affecting the assessment of culpability rather than as an automatic ground for exoneration. The rule should require a serious impairment, resulting from the influence, of the person's capacity to understand the nature of the conduct, control their behaviour in light of its unlawfulness or perceive facts correctly. Complete impairment may exclude responsibility, while a substantial reduction may diminish culpability or punishment.

Second, existing provisions on instigation, aiding, fraud, threats and perpetration through another must be interpreted in a technologically neutral manner for intentional digital manipulation. A new offence should be created only where a clear gap in protection exists. Otherwise, symbolic criminalisation weakens the requirements of legality and certainty.

Third, clear duties of oversight should be established for deployers of high-risk conversational systems. These include human intervention when a signal of vulnerability arising from age or mental condition appears, termination of dangerous interaction, warnings comprehensible to the user, independent audits and recording of incidents.

A gross breach of such a duty with knowledge of the possibility of serious consequences may provide a basis for criminal negligence. Meaningful human control consists not merely of a button and a position, but of the ability to recognise an error, change a decision and possess sufficient time and authority to do so [21].

Fourth, providers must fulfil their duty to preserve evidence. If the model version, system prompts or logs disappear after a dangerous event, the court cannot fairly assess human will and organisational knowledge. Privacy and trade secrets may be safeguarded through procedural protections, but they must not become a right to conceal necessary evidence in criminal proceedings absolutely.

Fifth, prevention should take precedence over the subsequent allocation of culpability. Testing models against dangerous influence, restricting psychological microtargeting, designing systems that respect the user's decision making process and providing an easy means of exit should be the principal directions of the law. The need to examine manipulation in generative artificial intelligence at the design stage has been identified in academic literature as a distinct research agenda [18].

Conclusion

When artificial intelligence induces a person to commit a crime, culpability does not mechanically move from one place to another. It is reallocated according to each participant's understanding, purpose, creation of risk and capacity for control. The fact that a person performed the final act does not always make them the sole culpable party. The fact that an algorithm exerted influence does not always make the person a victim.

The proper question for criminal law is not the metaphysical question of whether free will exists. It is whether this person, in these circumstances, possessed the legally significant capacity to understand the reason for criminal conduct, resist it and act otherwise. If the answer is affirmative, culpability remains. If the algorithm seriously impaired that capacity through deception, exploitation of vulnerability or digital coercion, culpability is reduced or excluded. If a provider or platform intentionally created the risk or grossly breached a duty of oversight, its independent culpability must also be examined.

In this sense, the title referring to the death of free will is a warning. What is dying is not human agency, but the conception of a decision as a closed event occurring solely inside a person's head. In the algorithmic era, will is formed within relationships, and culpability is fairly assessed through a map of knowledge and control within those relationships. The criminal law of the future will be strong through its refusal to allow people to hide behind machines, rather than through punishing machines as humans.

References

1. Henz P. Ethical and Legal Responsibility for Artificial Intelligence. Discover Artificial Intelligence. 2021. Vol. 1. Art. 2. DOI 10.1007/s44163-021-00002-4.

2. Simmler M. Responsibility Gap or Responsibility Shift. The Attribution of Criminal Responsibility in Human Machine Interaction. Information, Communication and Society. 2024. Vol. 27. No. 6. P. 1142-1162. DOI 10.1080/1369118X.2023.2239895.

3. Logg J. M., Minson J. A., Moore D. A. Algorithm Appreciation. People Prefer Algorithmic to Human Judgment. Organizational Behavior and Human Decision Processes. 2019. Vol. 151. P. 90-103. DOI 10.1016/j.obhdp.2018.12.005.

4. Köbis N., Bonnefon J. F., Rahwan I. Bad Machines Corrupt Good Morals. Nature Human Behaviour. 2021. Vol. 5. P. 679-685. DOI 10.1038/s41562-021-01128-2.

5. Matz S. C., Teeny J. D., Vaid S. S., Peters H., Harari G. M., Cerf M. The Potential of Generative AI for Personalized Persuasion at Scale. Scientific Reports. 2024. Vol. 14. Art. 4692. DOI 10.1038/s41598-024-53755-0.

6. Salvi F., Ribeiro M. H., Gallotti R., West R. On the Conversational Persuasiveness of GPT 4. Nature Human Behaviour. 2025. Vol. 9. P. 1645-1653. DOI 10.1038/s41562-025-02194-6.

7. Bigenwald A., Chambon V. Criminal Responsibility and Neuroscience. No Revolution Yet. Frontiers in Psychology. 2019. Vol. 10. Art. 1406. DOI 10.3389/fpsyg.2019.01406.

8. Susser D., Roessler B., Nissenbaum H. Technology, Autonomy, and Manipulation. Internet Policy Review. 2019. Vol. 8. No. 2. DOI 10.14763/2019.2.1410.

9. Ienca M. On Artificial Intelligence and Manipulation. Topoi. 2023. Vol. 42. P. 833-842. DOI 10.1007/s11245-023-09940-3.

10. Renzo M. Manipulation and Liability to Defensive Harm. Philosophical Studies. 2021. Vol. 178. P. 3483-3501. DOI 10.1007/s11098-021-01610-7.

11. Panattoni B. Generative AI and Criminal Law. Cambridge Forum on AI. Law and Governance. 2025. Vol. 1. Art. e9. P. 1-14. DOI 10.1017/cfl.2024.9.

12. Greipl A. R. Data Driven Learning Systems and the Commission of International Crimes. Concerns for Criminal Responsibility. Journal of International Criminal Justice. 2023. Vol. 21. No. 5. P. 1097-1118. DOI 10.1093/jicj/mqad031.

13. Bonicalzi S., De Caro M., Giovanola B. Artificial Intelligence and Autonomy. On the Ethical Dimension of Recommender Systems. Topoi. 2023. Vol. 42. P. 819-832. DOI 10.1007/s11245-023-09922-5.

14. Lu W. Inevitable Challenges of Autonomy. Ethical Concerns in Personalized Algorithmic Decision Making. Humanities and Social Sciences Communications. 2024. Vol. 11. Art. 1321. DOI 10.1057/s41599-024-03864-y.

15. Klenk M. Ethics of Generative AI and Manipulation. A Design Oriented Research Agenda. Ethics and Information Technology. 2024. Vol. 26. Art. 9. DOI 10.1007/s10676-024-09745-x.

16. Zeng Y. et al. How Johnny Can Persuade LLMs to Jailbreak Them. Rethinking Persuasion to Challenge AI Safety by Humanizing LLMs. Proceedings of ACL 2024. P. 14322-14350. DOI 10.18653/v1/2024.acl-long.773.

17. Elish M. C. Moral Crumple Zones. Cautionary Tales in Human Robot Interaction. Engaging Science, Technology, and Society. 2019. Vol. 5. P. 40-60. DOI 10.17351/ests2019.260.

18. Bernáth L. Can Autonomous Agents Without Phenomenal Consciousness Be Morally Responsible. Philosophy and Technology. 2021. Vol. 34. P. 1363-1382. DOI 10.1007/s13347-021-00462-7.

19. Mamak K. AI Agency and Criminal Responsibility. A Category Mistake. Inquiry. 2026. Online first. DOI 10.1080/0020174X.2026.2720904.

20. Coeckelbergh M. Narrative Responsibility and Artificial Intelligence. AI and Society. 2023. Vol. 38. P. 2437-2450. DOI 10.1007/s00146-021-01375-x.

21. Alon Barkat S., Busuioc M. Human AI Interactions in Public Sector Decision Making. Automation Bias and Selective Adherence to Algorithmic Advice. Journal of Public Administration Research and Theory. 2023. Vol. 33. No. 1. P. 153-169. DOI 10.1093/jopart/muac007.

