SOCIAL ENGINEERING CRIMES USING ARTIFICIAL INTELLIGENCE: LIABILITY FOR PHISHING AND VOICE CLONING

Xalilov Saidakbar Muratovich
Master's student, Public Administration Law (70420106)
https://doi.org/10.5281/zenodo.20589614

Abstract. This article analyzes the legal aspects of social engineering crimes carried out using artificial intelligence technologies. The research focuses on two of the fastest-growing and most dangerous manifestations — automated phishing based on large language models (AI-powered phishing) and voice cloning. The author sequentially examines the technical mechanisms of these crimes, their socio-psychological methods of influence, and issues of criminal-legal qualification. Furthermore, the article provides a comparative analysis of liability mechanisms against these crimes in international legal practice and national legislation, and formulates practical recommendations for the legal system of the Republic of Uzbekistan.

Keywords: social engineering, artificial intelligence, phishing, voice cloning, deepfake audio, fraud, cybersecurity, criminal liability, deception, identity spoofing.

Introduction

In the field of modern cybersecurity, alongside technical attacks — unauthorized access to systems, exploitation of vulnerabilities — a type of crime exists that is becoming increasingly dangerous and widespread: social engineering. Social engineering is based not on manipulating a technical system, but on manipulating the person themselves — their trust, fear, curiosity, or carelessness. Because the primary tool of this criminal art is human psychology, it cannot be blocked by software protection or a firewall.

Artificial intelligence technologies have increased the effectiveness and scale of social engineering to an unprecedented degree. Using large language models (LLMs), a criminal can now automatically generate millions of individualized, seemingly credible phishing messages. Voice cloning technology, in turn, allows a person's voice to be reconstructed from just a few seconds of recording and used in real time. When these two technologies are combined, a completely fabricated situation is created that gives the impression of communicating with "the person one trusts most," yet is entirely fake.

The purpose of this article is to comprehensively analyze the technical nature, socio-psychological mechanisms, and criminal-legal character of phishing and voice cloning crimes carried out with the assistance of artificial intelligence,

and to formulate recommendations on regulating liability issues for the legal system of the Republic of Uzbekistan.

The term social engineering — from the English "social engineering" — was first used in the field of information security in the late 1970s. In a broad sense, it is the art of psychological influence aimed at obtaining confidential information or material assets by manipulating individuals. In a narrow sense, it is a set of criminal techniques designed to exploit the human factor to gain access to computer systems or networks.

Among the classic types of social engineering are: phishing — prompting users via email or messages to click on fraudulent links and fill out forms; vishing — extortion via telephone calls; smishing — deception via SMS messages; pretexting — creating a false situation to gain trust; and baiting — setting a trap by exploiting curiosity or greed. All of these have reached a qualitatively new level of threat in the age of artificial intelligence.

Artificial intelligence has brought three fundamental innovations to social engineering. The first is scale: previously, a single criminal could deceive dozens of people; now it is possible to simultaneously influence millions of people, each in an individually tailored manner. The second is quality: messages created on the basis of LLMs are free of grammatical errors, strange styles, or suspicious signs, and can deceive even experienced individuals. The third is adaptability: based on the victim's profile, the AI changes its tactics in real time and selects the most effective method of influence.

Phishing — the act of prompting users through fake websites, emails, or messages to disclose their login credentials, passwords, bank card numbers, or other confidential information — remains one of the most frequently occurring types of attacks in cybersecurity. Classical phishing had a significant weakness: identical, non-personalized messages were quickly recognized by experienced users. For this reason, criminals developed spear phishing — that is, tailoring messages to a victim's specific situation by collecting prior information about them.

Artificial intelligence has fully automated spear phishing and brought it to a mass scale. Today, LLM systems automatically collect information about a victim from social networks, open databases, and professional profiles, and create a seemingly credible message for each person — tailored to their skills, professional activity, family situation, or current events. For example, a company employee may receive a fake message written in the exact style of their direct manager, containing the specific project name, a colleague's name, and the

current issue at hand. Distinguishing such a message from ordinary phishing is nearly impossible for most people.

Several new forms of AI phishing pose particular danger. In **Business Email Compromise (BEC)** attacks, a fake message is sent in the name of an executive or partner, demanding the immediate execution of a financial transfer. A modern LLM creates such a message in a completely convincing manner by imitating the target individual's working style, habitual phrases, and communication culture. In the **multi-stage phishing** scheme, the AI automatically conducts a multi-day "trust-building" process with the victim before proceeding to cause the primary harm. In this case, the victim involuntarily feels acquainted with the interlocutor and dismisses warning signs.

From the perspective of the criminal-legal qualification of phishing, the main problem is not genre confusion but identifying the responsible subject. Because the LLM process that created the phishing message is automated, it is difficult for the prosecution to point to a specific human action. However, this difficulty can be overcome: the person who configured the LLM for their own criminal purposes, set it in motion, and benefited from the result clearly exists. That person's deliberate actions fully encompass the objective and subjective elements of classic fraud or computer crimes.

Voice cloning technology — the AI-assisted synthetic reconstruction of a specific person's voice — has advanced so rapidly in recent years that it is now possible to create a convincing clone from just 3 to 5 seconds of a voice recording. This technology, also known as deepfake audio, was initially applied in the music and film industry to restore the voices of deceased performers. However, it was quickly adopted for criminal purposes as well.

Voice cloning technology is a particularly powerful tool for social engineering because human vulnerability to a familiar voice is embedded at an intuitive level. A parent recognizes their child's voice, an employee recognizes their manager's voice, a consumer recognizes a bank representative's voice in an instant, and trust in that familiar voice arises spontaneously. Criminals exploit precisely this psychological mechanism.

There are several typical manifestations of crimes carried out via voice cloning. The **grandparent scam** — frightening elderly people using a cloned child's or grandchild's voice to pressure them into urgent money transfers — is the most damaging voice-cloning crime worldwide. In the **corporate vishing** scheme, a company executive's cloned voice calls the finance department demanding the immediate execution of a money transfer. A 2019 case in Europe,

where the finance director of an energy company was deceived by a cloned CEO's voice and transferred €220,000, is recorded as one of the first major precedents in the field. Cloned voices of public officials are also increasingly being used for **political disinformation** purposes, placing the issue of national security alongside criminal liability on the agenda.

The boundary between lawful and unlawful uses of voice cloning technology is also problematic. Recreating a specific person's voice without their consent and using it for commercial or criminal purposes constitutes a violation of rights relating to that person, including the right to privacy and the right to personal interests. However, in most countries' legislation, a norm specifically designating the unlawful use of voice cloning does not yet exist, and adaptation of existing legal norms is required.

The primary difficulty in legally qualifying crimes carried out through phishing and voice cloning lies in correctly mapping these acts onto existing criminal norms. In the Criminal Code of the Republic of Uzbekistan, fraud — the misappropriation of property through deception or abuse of trust — is considered the primary applicable norm. AI-assisted phishing and voice cloning may also be interpreted as methods that substantively cover this corpus delicti.

However, several important aspects require separate analysis. First, in phishing and voice cloning, the means of deception is not a physical act but a technological simulation — a fake voice or fake identity. This expands the concept of "deception": the distinction between simple lying and technological falsification must be accounted for when describing the elements of the crime.

From the perspective of the subjective element, criminal intent is clearly manifested: the person who created the phishing message or cloned voice is aware of the criminal nature of their actions and the harmful outcome. However, the liability of the platform or service provider that created and offered the technology is a separate matter. If a voice cloning platform knowingly provided services while aware that the technology was being used for unlawful purposes, its liability may also be considered.

From the perspective of causation, the paradox of victim free will arises: the victim voluntarily transferred money or provided information, thereby exercising their free will. However, this free will was formed on the basis of an artificially created false reality. In most legal systems, it is recognized that establishing trust through a fabricated situation does not break the causal connection: even though the victim acted freely, this action is a direct result of the criminal's deception.

In the practice of qualifying crimes, phishing and voice cloning acts are typically examined through the lens of fraud, crimes in the field of computer information, or with the application of specific cybercrime norms in aggravated form. Furthermore, voice cloning inherently contains an element of identity theft, which may also be qualified as a separate offense.

Identifying the responsible subject in phishing and voice cloning crimes is more complex than in classic fraud, because the technology chain involves many participants. Each link in the chain can serve as a potential source of liability, but their levels of responsibility differ significantly.

The first link in the chain is the **direct criminal** — the person who sent the phishing message or made the call via voice cloning. Their deliberate act clearly and fully gives rise to criminal liability. The second link is the **technology platform or service provider**. If the platform was aware that its service was being used for criminal purposes and failed to take appropriate measures, it may, in most legal systems, be subject to civil liability and, in some cases, criminal liability. The third link is the **person who provided the voice source**. Where cloning is carried out through voice theft, the victim is unaware; however, if a person consciously provided their voice for misuse, they too may bear responsibility.

The issue of platform liability is a new manifestation of the global debate on the responsibility of online intermediary service providers. Under the tradition of Section 230 of the US Communications Decency Act, platforms were not considered liable for user-generated content. However, in AI-generated content, the platform is not merely a passive transit channel but an active participant in the content creation process. This requires rethinking the traditional principle of platform immunity, and legislation in a number of countries has already begun to change in this direction.

In regulating phishing and voice cloning crimes, states around the world are applying different approaches, though common trends are also visible.

In the United States, phishing is subject to criminal liability under the classic wire fraud and computer fraud laws. Since voice cloning is a relatively new phenomenon, no specific federal norm currently exists. However, some states — such as Georgia and Texas — have adopted specific laws designating the unauthorized use of a person's image or voice through artificial intelligence as grounds for criminal liability or civil action. At the federal level, the FCC (Federal Communications Commission) issued a ruling in 2024 explicitly prohibiting robocalls made with AI-generated voices.

In the United Kingdom, the Fraud Act 2006 covers acts related to phishing and identity theft, while the Online Safety Act 2023 introduced new obligations concerning platform liability and deepfake content. Fraud committed through deepfake audio may be prosecuted under the Fraud Act's provisions on deception and false representation.

In the European Union, the NIS2 Directive and the AI Act together establish liability standards for cyberattacks and high-risk AI systems. The AI Act classifies large-scale real-time surveillance and cloning of faces and voices as "high-risk" or absolutely prohibited, which represents a significant restriction for both criminals and lawful users.

In China, specific regulations governing "deep synthesis" (deepfake) technology came into force in 2022, requiring mandatory labeling of AI-generated content — that is, indicating that the content was created by AI. This approach is based on the principle of helping victims protect themselves by making the technological basis of deception transparent.

In the investigation and proof of phishing and voice cloning crimes, a number of specific problems exist that require adapting traditional forensic methodology to new conditions.

The first problem is source identification. Identifying the true source and creator of AI-generated content is technically complex. Phishing messages are often sent through several proxy servers, VPNs, and anonymization tools. Identifying the software tool used for voice cloning and linking it to a specific person requires intensive forensic work.

The second problem is proving a cloned voice. Within an investigation, distinguishing a cloned voice from a genuine voice requires specialized voice-forensic expertise. Such expertise has not yet been sufficiently developed in the forensic institutions of the Republic of Uzbekistan, which creates practical difficulties in proving such crimes.

The third problem is the admissibility of digital evidence. For phishing messages, call recordings, and voice samples to be accepted as evidence in criminal proceedings, it must be proven that they were obtained unaltered and were not manipulated. The fourth problem is the transnational character: most phishing and voice cloning operations are managed from different states, which requires international legal assistance and cooperation.

In regulating social engineering crimes carried out with the assistance of AI in the Republic of Uzbekistan, several priority directions may be identified. At the level of substantive law, it is necessary to designate fraud committed through

phishing and voice cloning as an aggravating circumstance — that is, the use of AI tools, mass scale, and technological falsification must constitute liability-enhancing factors.

It is also necessary to establish as a separate criminal corpus delicti the unauthorized reconstruction of a person's voice or image through artificial intelligence and its use for unlawful purposes. This would ensure liability even in cases where fraud has not occurred — for example, when voice cloning is used for disinformation or defamation purposes. Supervisory requirements for strengthening digital identification and authentication tools are also necessary: technical protection standards against AI-assisted phishing and voice cloning attacks for financial institutions and public services may be established at the level of subordinate legislation.

From a procedural standpoint, developing voice-forensic expertise capacity, expanding the current standards for working with digital evidence to cover AI-generated content, and improving the technical preparedness of investigative bodies are required. From a preventive standpoint, promoting awareness among the population about the methods of AI-assisted social engineering, the signs by which to recognize them, and ways to protect oneself — in partnership among the state, civil society, and the private sector — is one of the most effective means of preventing such crimes.

Conclusion

Social engineering crimes carried out using artificial intelligence — phishing and voice cloning — represent one of the fastest-growing and most directly harmful types of threats in modern cybersecurity. This research shows that these crimes constitute a new, technologically amplified form of traditional fraud, whose criminal-legal essence is preserved, but whose scale, level of believability, and complexity of proof have changed qualitatively.

On the question of the responsible subject, the main conclusion is that the responsible person is always clearly identifiable — it is the individual who used the technology for criminal purposes. The liability of platforms and service providers is determined based on the degree to which the technology facilitated the crime and whether there was awareness of it.

International practice shows that the most effective approach consists of the harmonization of three directions: adapting existing criminal norms to the AI context and designating them as aggravating circumstances; updating platform liability standards for AI-generated content; and enhancing technical-forensic

capacity. Consistent action by the Republic of Uzbekistan in these three directions will ensure effective combating of social engineering crimes.

As a final conclusion, it must be emphasized that although artificial intelligence enables the exploitation of human psychological vulnerabilities, with the proper construction of legal and technical protection systems, it is possible to mount an adequate response to this threat. In this endeavor, human critical thinking and digital literacy also remain the most important tools of defense.

References:

1. Hadnagy, C. Social Engineering: The Science of Human Hacking. — 2nd ed. — Hoboken: Wiley, 2018. — 320 p.
2. Chesney, R., Citron, D. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security // California Law Review. — 2019. — Vol. 107, No. 6. — P. 1753–1820.
3. King, T. C., Aggarwal, N., Taddeo, M., Floridi, L. Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions // Science and Engineering Ethics. — 2020. — Vol. 26, No. 1. — P. 89–120.
4. Europol. Facing Reality? Law Enforcement and the Challenge of Deepfakes. — The Hague: Europol Innovation Lab, 2022. — 45 p.
5. FTC (Federal Trade Commission). Voice Cloning Challenges Report. — Washington D.C., 2023.
6. FCC (Federal Communications Commission). Rules Targeting AI-Generated Robocalls. — Washington D.C., 2024.
7. Ferrara, E. The History of Digital Spam // Communications of the ACM. — 2019. — Vol. 62, No. 8. — P. 82–91.
8. Shariff, A., Bonnefon, J. F., Rahwan, I. Psychological Roadblocks to the Adoption of Self-Driving Vehicles // Nature Human Behaviour. — 2017. — Vol. 1, No. 10.
9. Council of Europe. Convention on Cybercrime (Budapest Convention), ETS No. 185. — Budapest, 2001.
10. European Union. Artificial Intelligence Act, Regulation (EU) 2024/1689. — Brussels, 2024.
11. UK Online Safety Act 2023. — London: HMSO, 2023.
12. China Cyberspace Administration. Provisions on the Management of Deep Synthesis Internet Information Services. — Beijing, 2022.
13. Criminal Code of the Republic of Uzbekistan. — Tashkent: Adolat, 2024.
14. Decree of the President of the Republic of Uzbekistan "On Measures for the Development of Artificial Intelligence Technologies". — Tashkent, 2024