# Improvement of Bank Fraud Detection Systems Through Voice Cloning and Manipulation

## Examining Voice and Speech Artificial Intelligence Usage in the Banking Industry

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

By
Drake Ferri

December 13, 2024

Technical Team Members:

Rhea Agarwal
Padma Lim
Vishnu Lakshmanan
Fahima Mysha
Baani Kaur

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Kent Wayland, Department of Engineering and Society

Greg Gerling, Department of Systems Engineering

*How can the banking industry deliver customers protection and security in the age of rapid innovation of Artificial Intelligence?*

## Introduction

In an era of ever-changing advancement and adaptation of AI, the banking industry faces a significant challenge: how to ensure customer trust and system security against increasingly organized attempts at fraud (Biswas et. al, 2021). While the advancement and adaptation of AI have certainly provided good for society, often in the form of more efficient operations in customer service, the quickly evolving nature of the technology has opened the door for malicious activities and fraud to impact consumer trust in the banking industry. As society continues to integrate automation and AI into daily life, it becomes essential to draw clear boundaries to safeguard the safety and security of consumers in the banking industry.

Specifically, the banking industry has undergone a collective effort to automate certain aspects, whether it is through customer service, risk management, or fraud detection. As these processes move towards automation using AI, customers become more likely to be unaware and uneducated about the risks involved in the system (Araujo et. al., 2020). This lack of awareness leaves customers vulnerable to new, complex methods of fraud that exploit gaps in security, such as voice cloning. As a result, the rapidly evolving technology has opened the door for AI to be used by fraudsters to “steal the identity” of a customer in many ways. With enough of a recording of a person’s voice, these fraudsters can replicate the mannerisms of a person and gain access to sensitive material (Kassis & Hengartner, 2023). In an attempt to dive deeper into the complexities of the issue of bank fraud through voice cloning, my group will conduct a technical research project that aims to identify vulnerabilities in banks' Automated Speaker Verification (ASV) systems and understand how criminals exploit them. It will be complemented by an STS research project that explores the potential threats to safety and security posed by voice-cloning deepfakes within the AI-human banking ecosystem. In tandem, these projects will create an

understanding of the intricacies of voice cloning and testing, in addition to their impact on the banking industry and consumer trust.

## Improving Fraud Detection Systems Through Voice Cloning and Manipulation

*How can voice cloning technology be used to identify and mitigate vulnerabilities in Automatic Speaker Verification systems in the banking industry?*

Numerous challenges arise in balancing AI-driven automation in banking with the guaranteed protection of customers. The growing threat of voice-cloning deepfakes highlights the risks of identity theft and unauthorized access to sensitive information in the banking industry (Bateman, 2020). Constant advancements in AI require ASV systems to be frequently updated to protect customers against security breaches. For instance, Kassis & Hengartner (2023) demonstrate in their IEEE paper that vulnerabilities in current ASV systems can be exploited in ways such as speech synthesis and voice conversion, underscoring the need for improvement.

For the technical project, my capstone team will investigate how banks' ASV systems detect cloned voices and develop a database of our own cloned voices to help stakeholders identify variables that distinguish real voices from cloned ones. We will research the successes and failures of voice deepfakes against voice authentication systems to enhance our understanding of detection techniques and system capabilities (Liu et al., 2023). Additionally, our research will cover quantifiable aspects of voice and speech to discover the key components and variables that make a person’s voice unique.

Our project will proceed in three main phases:

Creating a Library of Cloned Voices: Using both commercial and open-source generative AI tools, we will create a library of cloned voices, varying key factors found through research

such as pitch, tone, speech patterns, time of recording, and cloning tools used. These voices, generated using recordings of our team members, will be tested against a major bank's ASV system. Calls will be made to a live agent using Voice over Internet Protocol (VoIP) technology to play the cloned voices, which will say a pre-determined script in line with a typical banking call. From this, results will be collected by the bank and then distributed to our team in the form of categorical data that indicates the likelihood of realness/liveness of each voice.

Iterative Voice Cloning and Human Benchmarking: We will refine our cloned voices to increase their success rates based on the ASV system's feedback. For instance, if it is determined that the longer voice trainings produce more successful clones, we will be able to focus on longer times while continuing to test for other factors such as volume, cloning tools, background noise, etc. These refined voices will then undergo qualitative human testing to evaluate their ability to deceive human listeners, providing a benchmark for ASV system performance. This experimental phase will play real and cloned voices in an order unknown to participants, to which they will score based on their perceived likelihood of realness.

Testing Against Open-Source ASV Systems: Using a GitHub Python script, we will test the refined voices against open-source ASV systems. These systems will provide probabilistic data on realness and liveness, offering quantifiable insights into the factors that make cloned voices successful, further deepening our understanding of the factors and differences that impact the success of cloned voices. While the origins of many of these systems are difficult to trace, they provide value in that they differ from the standard ASV systems used in practice, and can be juxtaposed with the results from the bank to understand further the strengths and weaknesses of a bank's ASV system.

This process will give us quantifiable results, allowing us to determine the ways or combination of factors in which cloned voices are able to be successful. With an analysis of these results and the corresponding adjustments to our cloned voices, my capstone group will gain expertise in the structural strengths and weaknesses of ASV systems as well as AI voice synthesis. Following this, we aim to give banks further insight into ways in which these criminals gain access to sensitive information, highlighting key access points that may need to be improved upon as voice deepfake capabilities advance.

## Voice and Speech Artificial Intelligence Usage in the Banking Industry

*How does the adaptation of auditory AI through voice cloning impact consumer trust and security within the banking industry?*

According to the US Government Accountability Office, a deepfake can be defined as "a video, photo, or audio recording that seems real but has been manipulated with artificial intelligence technologies" (2020). AI advancements have allowed for these deepfakes to progress to a level that can oftentimes be impossible for the human senses to discern if the recording is authentic or not. Perpetrators and creators of deepfakes are constantly looking for avenues to leverage the technology and open the door for opportunities for themselves. Previous work has been done to characterize the overall threat of deepfakes that springs from increased AI adoption in the financial sector. The US Department of Homeland Security attempted to address the threat of deepfakes, expanding upon the idea that deepfakes bring forth a problem of not allowing society to believe what they see, losing trust in resulting interactions in their daily life. It has been reported by DeepMedia, a company specializing in tools to detect synthetic media, that over 500,000 video and voice deepfakes were shared in 2023 alone (Ulmer & Tong, 2023). This rapid

growth and innovation in deepfake usage have resulted in fraud detection technology that is "behind in developing tools to identify fake content" pertaining to audio deepfakes (Jingnan, 2024). As a result, fraud detection organizations work hand in hand with banks to implement systems that properly identify and restrict criminal activity (Lyeonov et. al., 2024)

This can be seen in the banking industry, where banks have begun to adopt voice-based security measures, such as voice authentication, to improve customer experience efficiency and enhance the organization of customer services (Narang et. al, 2024). However, these measures introduce vulnerabilities to the banking system, increasing the likelihood of access points for fraudsters to commit malicious activities, such as accessing sensitive financial information (Lyeonov et. al., 2024). Consequently, customer trust in banking systems is jeopardized, necessitating a balance between automation and security (Araujo et. al, 2020).

*STS Framework*

For my STS research topic, I will research AI voice deepfakes in the banking industry. Within the sociotechnical system of AI deepfakes in banking, there exist many components and factors that shape its functionality. The security of customers immediately comes to the forefront, as without proper security of their assets and information, no customer would do business with the bank. These customers work directly with the technology itself, interacting in ways such as automated call responses and voice recognition software. Hence, I will utilize the Social Construction of Technology (SCOT) framework, created by Trevor Pinch and Wiebe Biker, through my studies of the subject. This theory explains that societal factors and human actions drive technological advancement (Johnson, 2005). In the context of voice fraud in banking, SCOT theory allows for highlighting the negotiation between trust and security among

stakeholders, such as developers of AI, consumers of the bank, banking institutions, lawmakers and regulators, and fraudsters. The theory establishes that each stakeholder interprets the technology of voice cloning differently (Johnson, 2005), ultimately creating tension between security and convenience. Further exploration using this theory will reveal the societal factors influencing each stakeholder's trust or distrust with the increased adaptation of auditory AI in the banking industry.

*Methods*

While investigations have been done on deepfakes as a whole, there is limited research on the fraud that arises from voice and audio deepfakes within the banking industry. To assess the scope of the problem, I will extract prior surveys and questionnaires, such as JD Power's, where they found that only 28% of bank consumers believe that AI will improve their financial lives (2024). This analysis will allow for a quantification of the problem, identifying key areas of consumer distrust where security guarantees and increased consumer awareness of risks are necessary. When placing the results of these surveys in a timeline, the progression of public opinion on AI will become clear, which will allow for insights into the active shaping of voice AI technologies due to society.

To further address the question of the impact and resulting interactions caused by voice cloning, I will leverage news and incident reports of past instances of identity fraud through voice cloning. By analyzing cases like the incident where a finance worker in Hong Kong was deceived into facilitating a $25 million scam due to a deepfake call impersonating the Chief Financial Officer, key similarities and differences can be identified (Chen & Magramo, 2024). These comparisons help reveal trends and provide insights into recurring issues and

vulnerabilities, leading to a deeper understanding of such malicious activities. Paired with these reports, I will examine deeper through corresponding public statements by the affected company as well as public opinion of the instances to illustrate how they impacted consumer trust.

Although governments and regulatory bodies are often behind technological advancements, it is necessary to gather reports and conduct analysis on these attempts at combatting auditory fraud in the financial market. With reports and transcripts of events such as the US Congress Senate Committee's Hearing on AI and Human Rights, it can be seen whether current laws adequately address deepfake risks, and how holes in these regulations have impacted the advancement of voice AI usage in banking (2023). Further, regarding instances of voice cloning fraud, the resulting penalty or legal enforcement taken will help to understand how regulations shape consumer trust in the banking industry. Through the use of a SCOT analysis, this shaping of trust will reveal the impacts it has on previous and prospective auditory AI adaptations in banking.

Together, the information and evidence gathered will allow for the establishment of a comprehensive understanding of the impact that AI voice cloning and its implications have on consumer trust in banking.

**Conclusion**

By examining both the societal and technical dimensions of voice cloning technology in banking, my research bridges the gap between understanding emerging vulnerabilities and addressing them directly. Through the STS research, I aim to uncover how societal perceptions–shaped by trust, regulation, and stakeholder negotiations–impact the adoption and security of auditory AI in banking. This work will reveal critical factors influencing consumer

confidence and illuminate how voice cloning technologies challenge the relationships between banking institutions and their customers. The technical project will build on this understanding by exploring how voice cloning technology can expose weaknesses in ASV systems and, ultimately, strengthen them. Through the identification of vulnerabilities and development of targeted mitigation strategies, this work addresses an immediate security challenge banks face as they adopt voice-based systems.

Together, these projects highlight the interplay between technological innovation and societal trust, offering insight into how banks can balance efficiency and security in ways that resonate with consumers. More broadly, this research contributes to the larger conversation about responsible AI adoption in critical systems, where maintaining trust is essential. Looking ahead, this work sets the stage for further research on how rapidly evolving AI technologies impact functionalities dependent on identity verification while also raising questions about how the banking industry's regulatory and technological responses must adapt to stay ahead of malicious actors. By addressing the immediate concerns of voice cloning fraud in banking and analyzing its societal state and implications, this research will provide a basis for developing more resilient, AI-integrated systems that prioritize both security and public trust.

## References


Araujo, T., Helberger, N., & Kruikemeier, S. (2020). In AI we trust? perceptions about automated decision-making by artificial intelligence. *AI & Society, 35*(3), 611-623.

Bateman, J. (2020, July). *Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios,* Carnegie Endowment for International Peace. https://carnegieendowment.org/2020/07/08/deepfakes-and-synthetic-media-in-financial-system-assessing-threat-scenarios-pub-82237

Chen, H., & Magramo, K. (2024, February 4). *Finance worker pays out $25 million after video call with Deepfake "chief financial officer."* CNN. https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html

J.D. Power. (2024, February 26). *As new technology is integrated into the financial services industry, Most bank customers in the United States express distrust for AI.* J.D. power. https://www.jdpower.com/business/resources/new-technology-integrated-financial-services-industry-most-bank-customers-united

Jingnan, H. (2024, April 5). *Using AI to detect AI-generated deepfakes can work for audio - but not always*. NPR. https://www.npr.org/2024/04/05/1241446778/deepfake-audio-detection

Johnson, D. (2005). Social construction of technology. In C. Mitcham (Ed.), Encyclopedia of Science, Technology, and Ethics (Vol. 4, pp. 1791–1795). Macmillan Reference.

Kassis, A. & Hengartner, U. (2023, May 22-24) *Breaking Security-Critical Voice Authentication* [Paper Presentation]. 2023 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, United States. 10.1109/SP46215.2023.10179374

Liu, Xin & Tan, Yuan & Hai, Xuan & Yu, Qingchen & Zhou, Qingguo. (2023, October 9-12). *Hidden-in-Wave: A Novel Idea to Camouflage AI-Synthesized Voices Based on Speaker-Irrelative Features* [Paper Presentation]. 2023 IEEE International Symposium on Software Reliability Engineering, Florence, Italy. 10.1109/ISSRE59848.2023.00029.

Lyon, B., Tora, M., & O'Reilly Online Learning: Academic/Public Library Edition (2023). Exploring Deepfakes: Deploy Powerful AI Techniques for Face Replacement and More With This Comprehensive Guide. S.l.: Packt Publishing Ltd.

Lyeonov, S., Draskovic, V., Kubaščikova, Z., & Fenyves, V. (2024). Artificial Intelligence and Machine Learning in Combating Illegal Financial Operations: Bibliometric Analysis. *Human Technology*, *20*(2), 325–360. https://doi.org/10.14254/1795-6889.2024.20-2.5

Narang, Ashima & Vashisht, Priyanka & Bhaskar, Shalini. (2024). Artificial Intelligence in Banking and Finance. International Journal of Innovative Research in Computer Science and Technology. 12. 130-134. 10.55524/ijircst.2024.12.2.23.

Ulmer, A, & Tong, A. (2023, May). Deepfaking it: America's 2024 Election Collides with AI Boom. https://www.reuters.com/world/us/deepfaking-it-americas-2024-election-collides-with-ai-boom-2023-05-30/

United States Congress Senate Committee on the Judiciary Subcommittee on Human Rights and the Law, & United States Congress Senate (2024). Artificial Intelligence and Human Rights: Hearing Before the Subcommittee on Human Rights and the Law of the Committee on the Judiciary, United States Senate, One Hundred Eighteenth Congress, First Session, June 13, 2023. Washington: U.S. Government Publishing Office.