VIRTUAL ME BLOCKCHAIN-BASED SYSTEM FOR VIRTUAL RIGHTS OWNERSHIP

Aljawharah Alhammad, Aljoharah Alsurayyi, Reema Alshehri, Saba Alhoshan, Maali AlAbdulhafith

College of Computer Sciences and Information, Princess Nourah Bint Abdulrahman University, Riyadh City, Saudi Arabia

ABSTRACT

In the virtual realm, human rights face vulnerability, particularly with intangible rights such as the right to own the voice, which lacks tangible representation in the physical world. The increasing use of artificial intelligence (AI) intensifies the challenge of protecting virtual rights, as there is currently no established legal or technical defense against violations, especially concerning voice ownership. Our proposed solution employs blockchain technology and smart contracts, forming the Virtual Me system. This innovative system attributes the original voice to its rightful owner, mitigating violations and unauthorized usage. What sets our solution apart is its pioneering role in providing a technical foundation for safeguarding virtual human rights. The system aligns with the requirements of virtual human rights ownership, ensuring comprehensive protection and registration for their original owners.

KEYWORDS

Voice Property, Blockchain, Smart Contract, Voice Cloning, Deepfakes.

1. INTRODUCTION

Humans have many rights that cannot be represented in the physical world, such as the right of voice, personal identity and more. Which are known as virtual rights. Undoubtedly, these rights must be protected from unauthorized uses and violations, just like any other human rights. With the fast growth of technology and the use of AI, which is a technology that uses intelligent techniques to perform specific tasks, which made human virtual rights more susceptible to unauthorized and unlawful uses and violations. Currently, the world is moving towards using technological solutions and living a virtual lifestyle through interactions on various technological platforms. With the emergence of AI, which has gained significant popularity, it has greatly facilitated access to various domains. On December 16th, 2020, within four months of its release, voice modifications using AI tools were adopted by 2,225 out of 4,021 creators, representing 55.3% of the total [1]. This high usage rate indicates the rapid and widespread use of this AI tool in a very short period. As a result, new concepts and principles for humans have emerged that lack foundations or legislation to protect them from violations and unauthorized uses, such as human voice ownership.

Each individual owns a unique and distinct voice, but they do not have the right of its ownership since it is considered a virtual right. This lack of ownership rights leaves individuals unable to defend against violations and unauthorized uses. With the widespread use of AI technologies that can modify any human voice to read any text, these technologies stand out for their high accuracy

David C. Wyld et al. (Eds): NBIoT, MLCL, NMCO, ARIN, CSITA, ISPR, NATAP-2024
pp. 29-37, 2024. CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.140503

and fidelity in reproducing human voices, achieving a high level of similarity between the synthesized voice and other individuals' voices. Consequently, an experiment was conducted at Johns Hopkins University, where a group of individuals was exposed to a set of fake voices. Only 57% of the fake voices were correctly identified as fake compared to real voices [2]. This highlights the danger of accuracy in voice cloning when it comes to content published on the internet. Moreover, these technologies extend to threats, profit-seeking activities, legal issues, and other many problems. And, as voice cloning and synthesis have reached the point of fabricating and disseminating speeches in the names of presidents and scientists, producing songs using the voices of singers, and publishing them. In a recent incident in October 2023, Meta launched accounts on the Instagram platform for fictional characters with different names, benefiting from the voices and faces of influencers and celebrities through AI. This enables people to interact and communicate directly with these virtual personalities [3]. This leads to unhealthy obsession, idolization of celebrities, and significant manipulation of future generations. These posts have raised significant concerns among the public, as they fear violations that are not adequately defended, making voice ownership right a form of property that needs official protection [4].

Despite the significance impact due to the problem of unauthorized voice usage, there have been no strict legislation or solutions to safeguard the ownership rights of human voices used for cloning and synthesis through AI. Although some countries, such as China, have attempted to solve the spread of manipulated voices through AI by implementing regulations on social media platforms to protect individuals, but they have not specifically protected the voice itself. It has been found that such regulation is ineffective as they have been unable to control the distribution of fake recordings locally or globally [5] [6]. Virtual rights of humans vary and expand with the advancement of technology and AI, and one of the most affected rights in our current era is the right of voice ownership.

2. PROBLEM STATEMENT & SIGNIFICANCE

The problem statement & significance of these issues becomes evident as we examine three key problems. Firstly, there is the existence of unsecured virtual human properties, leaving individuals defenseless when these properties are violated. The second problem is the challenge of limiting and controlling individuals who manipulate voices and publish them on the internet particularly with AI's technology increasing accessibility and usability. Lastly, the absence of clear legal protection for individual’s voice ownership rights in AI-modified voices is deeply concerning. Consequently, we have come to realize a huge gap between the rapid advancement of technology using AI and the preservation of human virtual rights within these technologies.

Firstly, the existence of unsecured virtual human properties, such as voice ownership rights, a right that cannot be owned in the physical world, leaves individuals defenseless against their violation. Making it vulnerable when subjected to impersonation and infringement. This problem has spread globally, as virtual rights emerged alongside modern technologies. In February 2020, the ruling political party in India utilized AI-generated deepfake technology to manipulate campaign footage, showcasing a member of the legislative council speaking in the Haryanvi dialect instead of the original English. The opposing group manipulated the audio to make it appear as if the member was speaking in their regional accent [7]. This case demonstrates how individuals can easily manipulate AI technologies to portray political figures as supporters of their causes. Currently, there is no clear law or punishment to deter such actions. However, humans have not yet been able to officially register or prove the voices to their original owners, preventing them from safeguarding their rights in such instances and holding the perpetrators accountable, whether the exploitation of their voice is for profit or non-profit purposes.

Secondly, it is extremely difficult to control individuals who modify and distribute voices globally over the internet. As AI tools are also characterized by their accuracy, as research has shown that a short duration of any human voice (about 20 minute) is sufficient to create a readable speech using that voice, applying the original speaker's accent with a 99% accuracy rate [8].This has made many presidents, scientists, singers, religious scholars, and even individuals in danger and vulnerable to identity impersonation. In 2020 voice cloning technology was used to impersonate a German CEO and they succeeded in deceiving His British subordinate sent an electronic transfer worth $243,000 [9]. This indicates that people's beliefs in manipulated content is not limited to the present time but also extends to the beliefs of future generations due to the lack of verification of the authenticity of the released content and the difficulty of discernment. The ability to manipulate any human voice, including the voices of religious scholars, compromises the beliefs of future generations. Currently, AI technology is advancing rapidly to the point where it is becoming beyond human control. It has become a clear threat to religions and the accuracy of transmitting faith-related beliefs as well. In August 2023, many audio clips have been circulating on the TikTok platform under the names of religious scholars discussing controversial religious matters. This raised concerns among a percentage of viewers, as they face difficulties in distinguishing between the manipulated voice and the accuracy of the spoken content. Some individuals tried to raise awareness about this matter by reproduced the voices of deceased scholars and shared them on the TikTok application, warning against believing all published content [10]. but the characteristics of AI have led to a widespread use worldwide, thus making the ownership of human voice a problem that must be solved.

Thirdly, it is important to acknowledge that so far, no effective technology or legal system has appeared to address and solve the problems linked with fake voices. Modifying and disseminating voices remains legally unaccountable for individuals up to the present time. Therefore, there is an urgent need for intervention and the development of a comprehensive solution to address this issue. This situation highlights a significant gap between the rapid advancement of technology using AI techniques and the protection of human voice rights within these technologies.

3. PROPOSED SOLUTION

The Virtual Me platform proposes an innovative and unique solution to address the problem of individuals not owning their voices as a tangible asset in the physical world, which currently allows for global manipulation due to the absence of voice ownership documentation. Virtual Me platform utilizes blockchain technology to register human voice with high precision directly from the original person by reading a predefined text provided by us. This ensures the clarity of all necessary letter sounds and accent that prove voice ownership when needed for analysis and verification. The audio recording is encrypted and securely stored in the digital Virtual Me wallet. Each voice is encrypted with a unique code, creating a distinct voice fingerprint that guarantees the preservation of virtual rights and prevents unauthorized usage. This allows each user the right to record only one voice, which is then stored in the blockchain wallet for smart contracts.

The blockchain technology guarantees that the data is stored securely and cannot be breached or tampered with. Virtual Me platform provides direct ownership rights of human voice to its owner, like individuals owning their images in personal identification documents. Providing a unique voice fingerprint to each person like face id or any other common prints, this virtual ownership can be used for self-defense to proof the ownership in case of violation and many other uses. The platform targets anyone in need of safeguarding their rights from unlawful usage, both for individuals in general and professionals in the voice industry. Virtual Me platform enables the provision of a unique code that proves the legality of manipulated voices, allowing the original voice owner to legally defend themselves or provide authorization for commercial purposes such as advertisements, voiceovers, and other uses. Establishing voice ownership as a tangible digital

asset allows for accountability in case of violations. It also grants the voice owner the right to sell their unique telling approach or their voice to individuals in the voice industry.

By leveraging the Virtual Me smart contract blockchain-based platform, assigning one's voice can effectively mitigate the potential for voice manipulation and unauthorized dissemination across the internet without accountability or legal repercussions. The platform's robust technological infrastructure ensures that voice recordings are securely stored and linked to a unique digital identity. This immutable association establishes a clear and traceable chain of ownership, deterring creators from illicitly altering voices or engaging in disruptive activities. Through the utilization of smart contracts, Virtual Me enforces contractual obligations and predefined rules governing the usage and modification of registered voices. This fosters a sense of responsibility and legal compliance among users, as any unauthorized modifications or misuse can be readily identified and attributed to the accountable parties. The non-centralized nature and the transparency of the blockchain technology employed by Virtual Me provides an added layer of assurance. All transactions and modifications are recorded on the blockchain, creating an auditable and tamper-proof record. This enables efficient monitoring and detection of any illicit activities, empowering voice owners to take appropriate legal measures when necessary. By establishing a framework that aligns voice ownership with legal rights and responsibilities, Virtual Me facilitates a more secure and accountable environment for voice-related activities. It effectively discourages unauthorized manipulation and dissemination of voices, promoting ethical conduct and safeguarding the interests of voice owners.

4. RELATED WORK

Through our research, we found the most related technical-based, legal-based solutions, that closely resemble our innovative solution. First Technical-based solutions for addressing the proliferation of fake digital content, including deepfakes, encompass various approaches. One proposed solution involves the utilization of blockchain and smart contracts. Researchers Haya Hasan and Khaled Salah suggest tracing content back to trusted sources by leveraging the Ethereum blockchain and smart contracts. This system records the history of transactions related to digital content, allowing the origin of the multimedia to be traced. Additionally, it includes functions that enable secondary publishers to request permission for sharing, editing, and distribution based on predefined terms and conditions. However, there is a potential risk in terms of sensitive data being edited by secondary publishers [11].

Another technical solution is the Ethereum public blockchain system called *SelfKey*. This system enables individuals and corporations to regain ownership of their identity data and securely store it in a decentralized platform. It employs public and private key cryptography to restrict access to critical data and files. User identity documents are stored on their mobile devices, ensuring user control. The *SelfKey* system consists of the utility token KEY, the *SelfKey* wallet for storing personal information and accessing financial products and services, and the *SelfKey* marketplace for accessing various financial offerings [12] [13] [14].

Lastly in the field of music recognition, *Shazam* is a well-known system that accurately identifies songs based on their unique acoustic fingerprints. It utilizes the Fast Fourier Transform (FFT) as part of its spectrogram analysis. The FFT is a technique that efficiently calculates the Discrete Fourier Transform (DFT), converting a signal from the time domain to the frequency domain. Shazam divides the audio signal into short overlapping segments, applies the FFT to each segment, and obtains its frequency spectrum. This allows the extraction of the features and the creation of a spectrogram, forming the basis for fingerprinting and identification in the Shazam algorithm. The FFT plays a crucial role in *Shazam's* signal processing pipeline, enabling accurate frequency analysis of audio samples [15] [16].

The other solution was Legal-based solutions to address deepfakes encompass various laws introduced by different jurisdictions. The U.S. Congress introduced the *Deep Fake Act*, which requires any shared or intended to be shared "advanced technological false personation record" to include a disclaimer indicating that the content depicted is false. This law holds individuals accountable for sharing harmful deepfakes without proper disclosures, and intentional ignorance or removal of disclosures can lead to criminal charges. However, legal action under this law requires deepfakes to inflict considerable harm to individuals or society, and individuals cannot file lawsuits if the content includes a disclaimer stating it is fake [17] [18].

California enacted the Deceptive Audiovisual Media law, focusing on electoral considerations. This law addresses intentionally altered audio or visual media of a candidate within sixty days preceding an election, where such alterations may appear genuinely misleading. The law includes exceptions for sarcasm or parody but does not provide extensive protection for private individuals who may fall victim to deepfake manipulations [7] [19].

China's State Internet Information Office released the Provisions on the Administration of Deep Synthesis Internet Information Services, which broadly bans the use of AI-generated content. These provisions aim to increase the government's control over the Internet to prevent social and political disruption. They emphasize cybersecurity, real name verification of users, data management, marking of synthetic content, and dispelling rumors. The Chinese government relies on tech companies to enforce these regulations, with industry cooperation being crucial. The difficulties in controlling the spread of deep synthesized content lie in labels being created and preserved, as well as the ability of content to be decoupled from the service and spread independently. Once information spreads on the Internet, it becomes challenging to erase completely [5] [6].

5. PROPOSED SYSTEM & SIMILAR SYSTEMS COMPARISON

Building upon the previously discussed solutions, Virtual Me is introduced as an innovative system leveraging blockchain technology and smart contracts to create voice fingerprints to safeguard virtual rights, prevent unauthorized access, and ensure data integrity. Virtual Me is important for those who prioritize the authenticity of their voices and consider it an integral part of their work. To access the platform securely, users authenticate themselves through virtual me account authentication process after they sign with their MetaMask wallet, which is a crucial step for obtaining a unique voice fingerprint. Each voice fingerprint is assigned to a unique Content Identifier (CID) code based on its content, which will be generated by InterPlanetary File System (IPFS), then it will be stored in the blockchain via smart contract, ensuring resistance to modification or manipulation, since any alteration will produce an entirely different code. The blockchain's decentralized and distributed nature, combined with mandatory transaction approval and validation, establishes it as a robust choice for secure information storage. Account creation or login triggers the issuance of one-time password (OTP) to authenticate users. The system captures and stores user activities, personal information, authentication activities, and transaction details in MongoDB, which will serve our platform as a backend database. As the digital world evolves, Virtual Me will shape the future of virtual rights and digital asset ownership.

Table 1 Similar Systems Comparison

Features	Our Proposed System	Blockchain and Smart Contract	Ethereum Public Blockchain	Music Recognition System
Using blockchain	✓	✓	✓	×
Voice analysis (FFT)	✓	×	×	✓
Registering the Ownership of Virtual Rights	✓	×	✓	×
Decentralized storage using IPFS	✓	✓	×	×
Universally recognize hash code (voice fingerprints)	✓	×	×	×

6. IMPLEMENTATION PLAN

First, the system will require from the user to sign with their MetaMask wallet, by doing so users will gain access to Virtual Me platform. Then, the system will allow the user to create an account by collecting necessary information such as name, ID, password, and phone number. This information will be securely stored in the MongoDB database.

To authenticate themselves, the user will provide an OTP code, which will be verified by the system. The system will maintain logs and generate reports to keep track of authentication, verification, and identification activities. To request a voice fingerprint, the user will have the ability to record their voice using a microphone. The system will then convert the voice recording into a digital audio file and store it in InterPlanetary (IPFS) File System. The audio file will be hashed using IPFS hashing algorithms, generating a unique CID representing the hash of the content. This CID will be included in a transaction that will be stored in the blockchain through Solidity-based smart contract. The smart contract will store the CID and relevant user information in the blockchain, enabling its retrieval when needed.

Furthermore, the system will allow user to request deletion of their voice fingerprint. The user will indicate the purpose of their request, and the system will validate the deletion request, if validated, voice records and voice fingerprint will be deleted. Additionally, the system will have data backup and recovery procedures to maintain data integrity and facilitate recovery in case of system failures or disasters. It will also prioritize accuracy, aiming for a high level of precision in distinguishing between different voice fingerprints. In terms of design, the system will be designed to be compatible with various platforms and devices, allowing users to access and interact with it from different environments. The user interface (UI) will be straightforward and user-friendly, promoting ease of use, learnability, and navigation to enhance the overall user experience.

7. SYSTEM DESIGN

7.1. System Architecture

System Architecture acts as a conceptual plan that defines the structure, behaviour, and different viewpoints of a system. In the context of the Virtual Me system, internet connectivity plays a crucial role in facilitating user interactions since the system operates online. The MetaMask Wallet functions as the interface through which users can engage with the system, while MongoDB serves as the storage system for the required user data. Additionally, the User Interface supports user authentication and facilitates interaction with smart contracts (see figure 1).

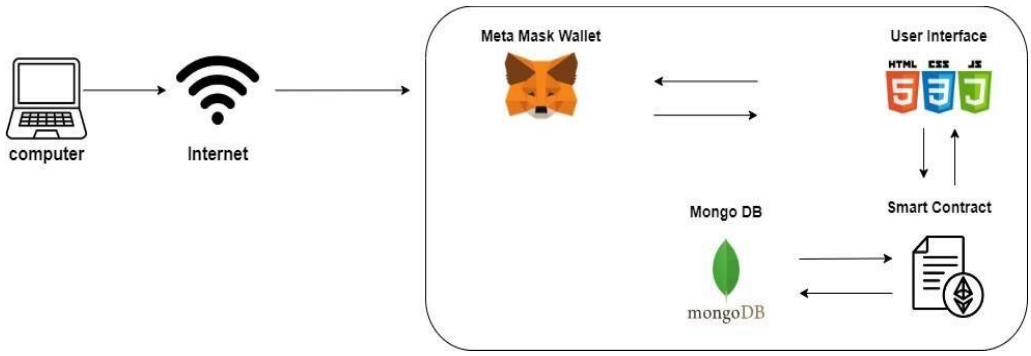

Figure 1: System Architecture of Virtual Me System.

7.2. System Architecture Layers Description

System Architecture is the conceptual framework that outlines the arrangement, functionality, and various perspectives of a system. In our application design, we have implemented a three-layered approach (see figure 2).

First, the front-end layer includes client-side scripts (HTML, CSS, JS) and the user interface, allowing users to interact with the system without relying on a conventional web server. Users can directly access the application without having to initiate a network connection with an external server. Second, the blockchain integration layer is crucial for connecting the application directly to the blockchain network without a centralized server. Smart Contract access facilitates interactions with the deployed smart contract, enabling secure storage and retrieval of voice fingerprints. MetaMask Integration allows users to interact with the blockchain securely and manage their wallets without relying on a central server. The MongoDB database continues to store user-related data and ensure data integrity and security.

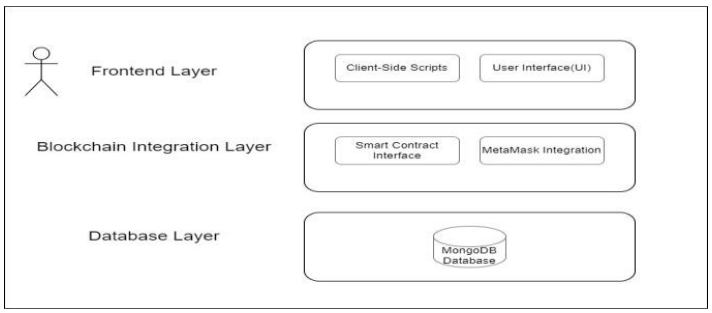

Figure 2: System Architecture Layers Description of Virtual Me System.

8. CONCLUSION

In conclusion, the fast growth of AI technology has made virtual rights, specifically the voice right, more vulnerable to unauthorized use and violations. The widespread availability of voice modification AI tools and the challenges in distinguishing between real and fake voices underscore the urgent need for legislative and technical measures to protect these rights. Unfortunately, there is currently a lack of strict legal-based and technical-based solutions to safeguard the ownership rights of human voices in AI cloning and synthesis.

To address this pressing issue, it is crucial to establish legal frameworks and innovative solutions that can effectively protect virtual rights, including voice ownership. The Virtual Me platform offers such a solution by utilizing blockchain technology. Through this platform, individuals can register and secure ownership of their unique voices as tangible assets. The voice is recorded directly from the original person, creating a distinct voice fingerprint that is encrypted and stored securely in a digital wallet. Each voice is assigned to a unique code, ensuring its authenticity, and preventing unauthorized usage. Virtual Me provides individuals with legal proof of voice ownership, similar to personal identification documents. This ownership can be utilized for self-defence against violations and allows individuals to authorize commercial uses of their voices. The platform caters to both individuals and professionals in the voice industry, offering them a means to protect their rights and establish accountability.

By leveraging smart contracts and blockchain technology, Virtual Me ensures the integrity of voice recordings, enforces contractual obligations, and maintains an auditable record of transactions. This framework discourages voice manipulation and unauthorized dissemination, fostering ethical behaviour and protecting the interests of voice owners. To access the platform securely, users authenticate themselves through their MetaMask wallet, which is crucial for obtaining a unique voice fingerprint. Each voice fingerprint is assigned to a unique code CID based on its content, which is stored in the blockchain through smart contracts. The decentralized and distributed nature of the blockchain, coupled with the mandatory approval and validation of transactions, ensures secure information storage and resistance to tampering.

Virtual Me aims to shape the future of virtual rights and digital asset ownership by capturing and storing personal information, authentication and user activities, and transaction details in a secure MongoDB backend database. This comprehensive record of platform interactions establishes a foundation for accountability and ensures the protection of virtual rights in an evolving digital.

Looking ahead, in future Virtual Me plans to commence the next phase of its study, which involves implementing and testing our system. This phase will involve the practical implementation of the platform, which include the system development. Bringing together all the components and functionalities to create a robust and secure environment for individuals. The system will go through several testing stages to verify its effectiveness. Prioritizing compatibility testing, system will be tested on various devices with different operating systems to ensure the compatibility and the smooth functioning. Next, we will verify the accuracy and precision of the enhanced voice captured by the FFT algorithm and the voice fingerprint generated from IPFS. The voice fingerprint and the enhanced voice of the recorded sound will be compared with other systems that provide voice encryption capabilities. We will verify the other functionalities on the system by conducting different tests on users in various environments to achieve the best user experience. Lastly, Virtual Me system will contribute to the ongoing efforts of safeguarding virtual rights in an increasingly interconnected world.

ACKNOWLEDGMENT

We would like to thank Eng. Muhammad Alagil and Nora Alagil for their support and generous grant in sponsoring our participation in this conference (Grant No.1). Our gratitude also goes to Princess Nourah University for facilitating the process of participation.

REFERENCES

[1] X. Zhang, "HOW DOES AI-GENERATED VOICE AFFECT ONLINE VIDEO CREATION? EVIDENCE FROM TIKTOK," 20 April 2023.

[2] Matthew Groha, Ziv Epsteina, Chaz Firestoneb, and Rosalind Picarda, "Deepfake detection by human crowds, machines, and machine-informed crowds," *Proceedings of the National Academy of Sciences,* 25 November 2021.

[3] Billie, "Instagram AI generated celebrities accounts," Instagram, [Online]. Available: https://instagram.com/yoursisbillie?igshid=OGQ5ZDc2ODk2ZA==. [Accessed 09 December 2023].

[4] "Posts have raised significant concerns among the public," TikTok, [Online]. Available: https://www.tiktok.com/@haussdrama/video/7288416213359807774?_r=1&_t=8hNvT2axvDW.[Accessed 09 December 2023].

[5] L.Floridi , E.Hine, "New deepfake regulations in China are a tool for social stability, but at what cost?,"*Nature Machine Intelligence,* vol. 4, no. 7, pp. 608-610, 20 July 2022.

[6] "Provisions on the Administration of Deep Synthesis Internet Information Services," 01 February 2023. [Online]. Available: https://www.chinalawtranslate.com/en/deep-synthesis/.

[7] M.Feeney, "Deepfake Laws Risk Creating More Problems Than They Solve," 1 March 2021. [Online]. Available: https://rtp.fedsoc.org/paper/deepfake-laws-risk-creating-more-problems-than-they-solve/. [Accessed 09 October 2023].

[8] HTPlay, "AI Voice Cloning with Unparalleled Quality," [Online]. Available: https://play.ht/voice-cloning/. [Accessed 10 November 2023].

[9] J.BATEMAN, "Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios," 20 July 2020.

[10] N. Mubarak, "Audience warning against believing all published content," TikTok, [Online]. Available: https://vt.tiktok.com/ZSN5M8QQp/. [Accessed 09 December 2023].

[11] K. S. Haya R. Hasan, "Combating Deepfake Videos Using Blockchain and Smart Contracts," *IEEE Access,* 17 March 2019.

[12] Jayana Kaneriya, Hiren Patel, "A Comparative Survey on Blockchain Based Self Sovereign Identity System,"*International Conference on Intelligent Sustainable Systems (ICISS),* 03 12 2020.

[13] "What Is SelfKey & How Does It Work? Who Created KEY?," Kriptomat, 19 Januray 2023. [Online]. Available: https://kriptomat.io/cryptocurrencies/selfkey/what-is-selfkey/. [Accessed 22 Novmber 2023].

[14] SelfKey, "Self-Sovereign Identity for more Freedom and Privacy," The SelfKey Foundation, 29 October 2023. [Online]. Available: https://selfkey.org/. [Accessed 22 Novmber 2023].

[15] "How does Shazam work? Music Recognition Algorithms, Fingerprinting, and Processing: Toptal®," [Online]. Available: https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognitionToptal Engineering Blog. [Accessed 22 Novmber 2023].

[16] S. Team, "An Industrial-Strength Audio Search Algorithm," [Online]. Available: https://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf.[Accessed 01 November 2023].

[17] Z. Schapiro, "Deep Fakes Accountability Act: Overbroad and Ineffective," *Boston College Intellectual Property and Technology Forum,* 11 January 2023.

[18] Yvette D. Clarke (NY), "Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019," *GovInfo,* 12 June 2019.

[19] L.Vazquez, "Recommendations for Regulation of Deepfakes in the U.S.: Deepfake Laws Should Protect Everyone, Not Only Public Figures," *Epstein Becker Green,* 2021.