International Journal of Cybersecurity Intelligence & Cybercrime

Volume 6 | Issue 2 Article 2

8-30-2023

Victimization by Deepfake in the Metaverse: Building a Practical Management Framework

deepfake; deepfake detection; metaverse; interpersonal cybercrime; cyberRAT; eysenck’s theory of criminality; routine activities theory

Recommended Citation
Stavola, J. & Choi, K. (2023). Victimization by Deepfake in the Metaverse: Building a Practical Management Framework. *International Journal of Cybersecurity Intelligence & Cybercrime: 6(2)*, .
Available at: https://vc.bridgew.edu/ijcic/vol6/iss2/2

Stavola & Choi. (2023). *International Journal of Cybersecurity Intelligence and Cybercrime, 6*(2), 3-20.

Victimization by Deepfake in the Metaverse: Building a Practical Management Framework

Julia Stavola*, M.S., Boston University, U.S.A.
Kyung-Shick Choi, Ph.D., Boston University, U.S.A.

Keywords: deepfake; deepfake detection; metaverse; interpersonal cybercrime; cyberRAT; eysenck's theory of criminality; routine activities theory

Abstract:
Deepfake is digitally altered media aimed to deceive online users for political favor, monetary gain, extortion, and more. Deepfakes are the prevalent issues of impersonation, privacy, and fake news that cause substantial damage to individuals, groups, and organizations. The metaverse is an emerging 3-dimensional virtual platform led by AI and blockchain technology where users freely interact with each other. The purpose of this study is to identify the use of illicit deep fakes which can potentially contribute to cybercrime victimization in the metaverse. The data will be derived from expert interviews (n=8) and online open sources to design a framework and provide solutions to mitigate deepfake-related victimization in the metaverse. This study identifies and further suggests a framework for advocacy of deepfake crime victimization in the metaverse through the application of the routine activities theory, as well as offender motivation and potential explanation of criminal behavior through Eysenck's theory of criminality.

Introduction

The development of the internet over the past few decades has advanced the means of how we connect to others socially, retrieve our daily news, and survive financially, especially after the COVID-19 pandemic. A successful functioning society relies heavily on the use of technology through the web and as a result, has increased the risk of cybercrime victimization for its users. The internet has evolved from Web 1.0 to Web 3.0, to the most recent development, the metaverse. Previous literature predicts that by 2026, 25% of the internet population will spend approximately an hour a day in a metaverse (Galer, 2022 para 1.). In addition, a recent survey conducted by Pew Research Center and Elon University found that 54% of experts believe that the metaverse will become a well-functioning component of half of a billion Internet users globally by 2040 (Anderson, 2022, p.197). With the increasing use of socialization in the metaverse comes the increased risk of victimization of users to crimes such as deepfake-related interpersonal cybercrime. Interpersonal cybercrime is defined as any act of criminal activity in cyberspace by an offender against another individual through digital interaction and communication, whether they have a legitimate or imaginary relationship (Moshin, 2021, p.3). Interpersonal cybercrimes can include but are not limited to sexual violence, cyberstalking, cyber harassment, cyberbullying, cyber roasting, and gender-based cybercrime through technology-facilitated abuse and violence (Moshin, 2021, p.4). Previous literature suggests that interpersonal cybercrimes create psychological turmoil for victims, aiding in the development of fear, stress, anxiety, and depression, and have been recorded to even lead to death by suicide in many cases (Moshin, 2021, p.4).

*Corresponding author
Julia Stavola*, M.S., Department of Criminal Justice, Boston University, 1010 Commonwealth Ave, 5th Floor, Boston, MA, 02215, U.S.A.
Email: juliaa@bu.edu

Interpersonal crimes are already on the rise in the metaverse. According to CNBC, a psychotherapist who conducts research on the psychological and physiological impact of the metaverse, Nina Jane Patel (43) reported that she was verbally and sexually harassed by four male avatars in the metaverse, who then gang-raped her avatar and took photos of the crime. Patel highlighted that her physiological and psychological reactions were extremely realistic, making it feel like the abuse was occurring in the physical world (Das, 2022, para. 3). In an interview with CNBC-TV18, Patel claims that she aims to use her experience as a motive to make the metaverse a safer place for children, as they can become prime targets for cybercrime (Das, 2022, para. 9). Meta also responded to this issue, reporting to CNBC-TV-18 that they are introducing a new safety measure called ‘Personal Boundary’ for Horizon Worlds and Horizon Venues, which restricts avatars from entering an invading distance with another avatar (Das, 2022, para. 20).

Before the development of the metaverse, deepfake interpersonal crimes attacked the surface web. Raffaela Spone (50) was arrested for cyberbullying and charged with three misdemeanor counts of the third degree of cyber harassment of a child and harassment in March 2021 for posting manipulated media of three members of her daughter’s cheerleading squad on the web (Gamiz, 2023). Law enforcement reviewed the media and determined it to be a production of deepfake. Spone was court-ordered to complete a mental health evaluation to determine any relevant underlying psychopathological illness that instigated her criminal behavior (Katro, 2022 para.8). Additionally, a recent case of deepfake-related financial crime occurred this May (2023) in China (Reuters, 2023, para.1). It was found that the offender had used deepfake technology to face swap an image and impersonate a friend of the victim during a video call, further scamming the victim to pay them 4.3 million yuan, which approximates to $622.000 in U.S currency (Reuters, 2023, para.3). This can be predicted to occur in the metaverse as well, as deepfakes can be used to impersonate other individuals, or avatars, and commit similar financial fraud (Tariq, 2023, p.2).

The term metaverse is a compound name for the latest development of the World Wide Web, combining the prefix “meta” - meaning beyond, with the term “verse”, which is short for the universe (Bale et al., 2022 p.1). The metaverse is an interactive environment built on blockchain and internet technology that combines virtual reality, augmented reality, digital reality, and actual reality where people in the physical world will be able to interact with each other through avatars. (Moioli, 2022 para. 4). These avatars are 3D virtual people who are created to be ‘your digital twin’ and can mimic our real-life emotions, appearance, facial expressions, and personal characteristics using machine-learning techniques (Yang et al., 2023, p.1). Although the use of avatars is essential in navigating the world of virtual reality, it also increases the risk of deepfake crime as it targets digital humans, which includes avatars. Expert witness in computer forensics Jerry Bui (2022) suggests that metaverse avatars are at an equally high risk of being manipulated using deepfake as any other internet program is, simply because of their digital composition (Bui, 2022 para.7).

Deepfakes are the manipulated product of media using artificial intelligence and the machine learning techniques of deep learning, through deep neural networks with the aim to deceive its viewers. (Whittaker et al., 2020, p.92, Bui, 2022 para.3). As deepfake is an emerging cybercrime, the average internet user cannot detect synthetic media with the naked eye. A recent study found that out of 93 participants, 65% of them are totally unaware of deepfake, only 13% are confident that the general population of internet users will be able to detect deepfake, and 57% of participants believe that the average internet user will not be able to differentiate a deepfake from authentic media (Ahmed et al., 2021, p.3.). These findings especially support the need for mitigation of deepfake not only on the surface web but also as a future mitigation framework for deepfake in the metaverse, as internet users are subjected to the risks of becoming victims of online interpersonal cybercrimes as deepfakes are a critical tool in producing fake news, defamation, pornography, and more (Ahmed et al., 2021 p.1). Cyber-sexual crime and cyber harassment are also examples of interpersonal cybercrime that can result from the production of deepfakes on both the surface web and the metaverse, and as previous cases prove, cause detrimental psychological damage to victims. Additionally, deepfake can be used to commit interpersonal crimes such as identity theft and cloning of avatars (Tariq et al., 2023).

Metaverse founders are aware of deepfake and the risk that it poses to the platform (Castillo, 2022, para. 8). As deepfake has become a prevalent issue on the surface web, it is important to understand the patterns and impacts that it can have on the metaverse. Through the application of the routine activities theory, this study demonstrates the relationship between suitable targets of deepfake and interpersonal cybercrime victimization. These victims often react with strong psychological responses, even if these crimes are taking place only in the virtual world or the metaverse (Kirwan & Power, 2013, p.14). In addition, cybercriminals often are diagnosed with serious psychopathologies such as neurodevelopmental disorders, schizophrenia spectrum, and other psychotic disorders, bipolar disorders, depressive disorders, anxiety disorders, obsessive-compulsive disorders, neurocognitive disorders, as well as various personality disorders (Woods, 2022 p 102-129). Eysenck's theory of criminality suggests that these diagnoses lead to an increase in offending (Gudjonsson, 2016, p.106). To effectively mitigate cybercrime in the metaverse and decrease the risk of victimization, users need to be aware of the types of cybercriminals on the internet.

The purpose of this study is to identify the potential issues of cybercrime victimization in the metaverse, specifically focusing on deepfake-related interpersonal cybercrime and providing a mitigation framework as well as support to victims. Using the routine activities theory and Eysenck's theory of criminality as guidelines, this study will further explore the psychological victimization of deepfake-related interpersonal crime and provide a potential behavioral explanation of deepfake interpersonal crime offenders. Additionally, this study will therefore suggest mitigation solutions through analysis of previous literature's capable guardianship suggestions of the routine activities theory. The methods used will include analysis of previous literature, application of supported theories, and execution of expert testimony. These methods combined will furthermore assist in the development of policy, legal, awareness, and technological approaches to prevention.

Theoretical Applications

Motivated offender

Eysenck's theory of criminality (1964) argues that mental illness, specifically psychosis, leads to criminal behavior, and is measured using the Eysenck Personality Questionnaire (Gudjonsson, 2016, p.106). Eysenck (1977) also developed a criminal propensity scale to measure these antisocial behavioral components of adult criminals by identifying their levels of personality dimensions such as extraversion, neuroticism, and psychoticism (Saklofske et al., 1978, p.1). Further research supports this theory, suggesting that psychological factors are of central importance in relation to both the causes of crime and its control (Eysenck & Gudjonsson, 1989, p. 247). The psychological factors of criminality are related to an individual's genetic composition that combines with factors of criminal propensity to increase their likelihood of offending (Gudjonsson, 2016 p. 106). This theory broadly suggests that criminals and psychopaths display behavioral patterns through their personality disorders which are identified through a biological genetic spectrum and make it difficult for the individual to resist engaging in acts of deviance, leading to psychopathic and criminal behavior (Goldsmith, 1981, p.1). Lickiewicz (2011) further suggests that the cyber offender typically originates from a pathological family where a history of hereditary mental disease is present in their ancestry (Lickiewicz, 2011, p.241).

Previous literature indicates that approximately 25% of the world's population suffers from mental or neurological disorders; a potentially significant number of digital users could be represented in this group (Woods, 2022, p.95). Individuals who live with antisocial personality disorder are likely to be offenders of crime due to their deceitfulness and manipulation, poor social conformity, impulsivity, and aggressiveness, which could aid in the violent and non-violent behavior patterns of cyberbullying, online harassment, cyberstalking, and more (Woods, 2022, p.117). In the metaverse, the offenders behind these crimes are the individual behind the avatar, as Cheong (2022) highlights that the violence of one avatar to another in the metaverse could cause psychiatric harm to the physical individual. (Cheong, 2022, p.22).

In relation to Eysenck's theory of criminality, a motivated offender who lives with mental illness is often driven by their psychological need to engage in criminal behavior. Furthermore, the Durham test of New Hampshire defines criminal insanity as unlawful acts that are the direct product of mental disease, whereas crimes could be considered a direct product of nearly any symptom of mental illness (Peterson et al., 2014, p.2). This supports Eysenck's theory of criminality by encouraging that the relationship between mental illness and criminal behavior is significant, especially with interpersonal cybercrimes, which could include deepfake-facilitated crime.

Target suitability

In today's modern society, individuals utilize the surface web as an essential factor of life, subconsciously putting themselves at risk of being a victim of cybercrime. While online engagement on the surface web can have many benefits, it increases the risk of cybercrime victimization. Cohen and Felson (1979) suggest that an individual's lifestyle essentially makes them a suitable target, which includes vocational and leisure activities such as social interaction through online engagement (Choi, 2008, p.311). In summary, online engagement increases the risk of users becoming a victim of cybercrime, specifically five main types: social engineering and trickery, online harassment, identity-related crimes, hacking, and denial of services and information (Woods, 2022, p.99).

Choi's (2008) cyber-routine activities theory (CyberRAT) also acknowledges suggestions of characteristics that pose an internet user a high risk of being a suitable target for cybercrime (Choi, 2017, p.2). This theory suggests that risky behavior is a large part of what increases an individual's risk of becoming a victim of cybercrime, as they lack cybersecurity measures and therefore are more susceptible to cyber interpersonal violence victimizations such as cyber-harassment and cyber-impersonation (Choi, 2017, p.3). Since deepfake can be used to facilitate these crimes, this theory can also be used for development of this study's framework.

Felson (1998) suggests that there are four measures of becoming a suitable target of cybercrime: the value of the suitable target, the inertia of the target, the physical visibility, as well as the accessibility of the suitable target, the inertia of the target, the physical visibility, as well as the accessibility of the target (Choi, 2008, p.312). This can also be suggested for the metaverse. For example, Roblox, an up-and-coming metaverse platform that allows users to use games created by other users, targets, among other groups, tween (ages eight to 14) internet users who can potentially become victims of cybercrime by engaging with this platform. In 2021, it was reported that out of the over 12 million gamers on Roblox, 54% of gamers were under the age of 13, and about 90% of the game developers were over the age of 18 (Liao, 2022, para.3). In addition, Roblox's fastest growing demographics are individuals between the ages 17-24, with over 50% of users at least 13 years old (Liao, 2022, para.2). With the high population of young gamers, it can be concluded that children are at high risk of becoming victims of cybercrime due to the accessibility that adult gamers have to them. Pew Research Center (2022) predicts that users of the metaverse will become addictive, therefore increasing victimization as they are less aware of the risks (Anderson & Rainie, 2022, p.23). Therefore, suitable targets of cybercrime in the metaverse will have less of a sense of reality and will easily be manipulated by other metaverse users.

Capable guardianship

This study analyzes four main ways to apply capable guardianship towards the issue of deepfake-related interpersonal cybercrime on the surface web and in the metaverse. These include legal solutions, policy solutions, technical solutions, and awareness solutions.

Previous Literature on Legal Solutions

Previous literature on legal solutions suggests that responses to the production and distribution of deepfake for interpersonal cybercrime use both public and private law to enforce legality (Meskys et al., 2020, p.6). In addition, there are several bills and acts that have been recently introduced to address the issue of deepfake creation. For example, The Malicious Deep Fake Prohibition Act of 2018 (S. 3805), establishes that cybercriminals who create and distribute deepfake media on the internet are to be charged with a criminal offense (Ferraro, 2019, p.6). Additionally, the pending bill, An Act to Protect Against Deep Fakes Used to Facilitate Criminal or Torturous Conduct criminalizes the distribution of deepfakes on the internet that is used as an accessory to commit other criminal acts (Ferraro, 2019, p.12). These bills cover the issue of deepfake-related crime on the surface web, however, there is yet to be established a legal solution for these offenses in the metaverse. However, since the issue of interpersonal cybercrime in the metaverse is a rising problem, legal solutions are slowly being introduced to try and justly mitigate these crimes in the future. Additional sources suggest a potential action that virtual communities can take to ensure procedural fairness, which includes creating oversight mechanisms to monitor obedience to rules and punish those who break the laws (Suzor, 2010, p. 60).

In relation to Eysenck's Theory of criminality, criminal liability is to be determined based on the offender's mental wellness. Previous literature highlights the difficulty in assessing the mental state of the individual behind the avatar to further establish a sanction (Cheong, 2022, p.22). It is further suggested that a statutory remedy would begin with the pre-action discovery to determine the identity of the individual behind the avatar who is responsible for imposing tortious and psychiatric harm to the victim, and then proceed with legal proceedings in the real world (Cheong, 2022, p.24).

Interpersonal cybercrime in the metaverse deserves legal sanction for offenders who commit these crimes. Previous literature suggests that in response to the potential criminal activity in the metaverse of stalking, assault, abusive conduct, child pornography, kidnapping, intellectual property law infringements, and financial frauds, law enforcement of the metaverse should use well-established and successful procedures of the physical world (Kasiyanto & Kilinc, 2022, p.17). However, the results of a study on the seriousness of interpersonal cybercrime in comparison to physical interpersonal crime illustrated that only approximately two-thirds of the subjects considered these offenses to be equally relevant in seriousness (Leukfeldt & Holt, 2019, p.342). In contrast, 11.1%-25.8% of subjects responded that they believed cyber-harassment to be a less serious offense (Leukfeldt & Holt, 2019, p.342). In an additional study conducted by Broll & Huey (2015), police officers were interviewed on their perspective of the seriousness of cyberbullying. The results displayed that these officers acknowledged cyberbullying to be a non-criminal interpersonal incident (Leukfeldt & Holt, 2019, p.343). This emphasizes the need for law enforcement training to educate these officers on the seriousness of interpersonal cybercrime on the surface web as well as the metaverse.

Previous Literature on Policy Solutions

To successfully enhance the security and safety of any internet platform, public policy is essential. The 2021 AR/VR Policy Conference highlighted the importance of enhancing safety for the users of virtual reality, as their realities can become manipulated, which raises concerns for physical, emotional, and mental harm to the users (Dick, 2021, p.4). A key takeaway from this conference to address this issue was the plan for future policy procedures, and suggest the collaboration of policymakers, industry leaders, civil society, and users of the metaverse (Dick, 2021, p.1). Furthermore, Harrell (2022) suggests a policy solution to deepfake distribution in the virtual worlds which include the use of informal industry guidelines as well as formal guidelines of governmental legislation (Harrell, 2022, p.1).

VR authentication methods can play a role in mitigating crime in the metaverse amongst avatars, including information-based authentication, biometric authentication, and multi-model authentication (Kurtunluoglu, 2022, p.3). These three methods ensure that the user who logs on to VR through their headset is in fact the owner of that headset. This increases security by ensuring user identity, therefore decreasing the opportunity for deceiving other avatars with deepfake. In other words, if the user logging on does not match the user identity in virtual reality, they will not have access to the verse and therefore cannot interact with other users.

To control the appropriate social communication between individuals on the metaverse, policy solutions such as informal and formal social control mechanisms are equally important. Informal crime control mechanisms that are implemented in virtual realities by the community, such as community guidelines and self-regulations, are referred to by Boellstorff (2008) as ‘grassroot governance’ (Parti, 2011, p.662). Although community guidelines can be effective, there are many criticisms (Parti, 2011, p.662). Examples of such criticisms include biased motives by administrators and members of the community who hold power over these guidelines, the failure of vigilantes to intervene in ad-hoc individual actions, and the procedure of executing an abuse report may put the victim at a higher risk of becoming a target for future abuse (Parti, 2011, p.662). For these reasons, technological informal crime control mechanisms are suggested to be the most effective, as they are flexible to the community’s needs, unbiased, and provide an immediate response to intervening violence (Parti, 2011, p.662). However, there could be instances where the volume of violence demands a higher control mechanism, which is where formal crime controls are needed.

Previous literature suggests that strict and detailed formal controls such as legal regulation and the presence of federal authorities are needed, including formal crime control mechanisms that provide increased security through police surveillance and criminal prosecution, such as the Child Exploitation and Online Protection Centre (Parti, 2011, p.655, p.663). To incorporate both informal and formal controls on a virtual platform, Parti (2011) suggests creating simple community guidelines and legal regulations, incorporating traditional criminal laws that apply to real-life standards as well as cyberspace, and requiring law enforcement patrol teams who thoroughly understand the components and threats of a virtual community such as the metaverse (Parti, 2011, p.667).

In response to crime in cyberspace and virtual worlds, it is important to use cyber ethics as well as AI ethics to provide a framework for policy solutions in mitigating deepfake interpersonal cybercrime on the surface web and in the metaverse. As there becomes a rapid increase of deepfake technology, issues of AI technology are on the rise. In response to this issue, countries around the globe are developing AI ethical guidelines to address the legal and ethical problems of individuals online using AI to communicate (Han, 2022, p.4). The development of these guidelines can cause a major shift in the increase of mitigating cybercrime issues on all internet platforms, including the metaverse.

While implementing policy procedures in response to mentally ill users in cyberspace, Montasari et al., (2018) explain that the psychology of cyber-attacks should be heavily considered while developing a policy procedure to enhance cybersecurity (Montasari et al., 2018, p.1) Although the metaverse is recently developed, the policy has begun to acknowledge the mentally ill population that has already been shown to be prevalent in the users. Therefore, VR, AR, and MR are being used on the metaverse platform to assist with the diagnosis and treatment of mental illness in the metaverse (Usmani et al., 2022, p.1). However, there is no current mental health clinic available for users of the metaverse to utilize for victimization counseling.

Although there is not yet an established treatment program for offenders in the metaverse, it has been suggested that the use of virtual reality can support offender rehabilitation in hopes of reducing crime, especially the interpersonal cybercrime of sexual offending. As many sexual offenders in cyberspace struggle with mental illness, this treatment program through virtual reality could be extremely beneficial in providing cognitive insight from offenders' perspectives (Pettifer et al., 2022, p.46).

Previous Literature on Technical Solutions

Legal professionals suggest that the metaverse greatly increases risks of advanced deepfake technology, which furthermore increases the opportunities for users to represent other users illicitly through deepfake generation (Pepper, 2022, para.4). It is therefore important to discover effective technical solutions to mitigate these risks, as well as determine the effectiveness of applying the existing technological deepfake detection techniques to the upcoming issue of deep fake in the metaverse.

Prior research has been conducted on deepfake detection techniques, as well as which were effective and ineffective. By using these detection techniques, it will be possible to identify if a source of media is authentic or synthetic. For example, a previous study (2022) introduced a new forensic detection technique for deepfake creation with corneal reflection of deepfake media by presenting the image on a screen and displaying a distinct pattern to expose the synthetic manipulations (Guo et al., 2022, p.1). In other words, the detection technique will identify the probing pattern of the manipulated images by comparing the corneal reflection with the probing pattern (Guo et al., 2022, p.2). It is expected that this technique will be a satisfactory approach to detecting deepfake media in the VR based metaverse as well since it can also be used to expose unauthorized use of synthetic models (Guo et al., 2022, p.5).

Additionally, security measures in the metaverse should be taken to ensure the protection of users from interpersonal cyber-attacks. Although there are established detection techniques to increase the cybersecurity of deepfake creation and distribution, there are limited technological developments aimed to prevent cyber-attacks in the metaverse. A potential reasoning behind this is that the metaverse is a recently developed platform, therefore, the realms of cyber security techniques for interpersonal cybercrime in virtual reality have not yet been explored. For this reason, it is imperative to initiate a mitigation framework that would apply to the metaverse.

Previous Literature on Awareness Solutions

There is a lack of awareness programs that are currently established for cybersecurity of deepfake-related interpersonal cybercrime in the metaverse. However, there are attempts to raise awareness for individual crimes such as deepfake and interpersonal cybercrime, which can be applied to the increase of these crimes in the metaverse.

From a psychological perspective, the MINDSPACE framework developed by Coventry et al., (2014) has been used to set an example of developing a cyber security awareness program through psychology and behavioral insights (Stalans & Donner, 2018, p.297). By educating the public using the MINDSPACE framework, it can be predicted that cyber security will increase through a change in practice on the internet. While this framework is intended to be applied for cybercrime on the surface web, there is reason to suggest that it can also be applied to cyber security awareness for crime in the metaverse.

There are many awareness programs and health services that are provided to victims of such internet crimes. An example includes the Innocent Justice Foundation, funded by the Internet Crimes Against Children task force, which provides a comprehensive mental health program for those who are impacted by online child sexual abuse. There are also programs designed to aid individuals experiencing incarceration who suffer from mental illness, such as the Justice and Mental Health Collaboration Program, developed by the OJP Bureau of Justice Assistance (DOJ, 2009, p.1). The goal of this program is to "improve access to effective treatment for people with mental illnesses involved with the justice system. This contact can be through arrest, court appearances, community-based supervision, incarceration, or in the community following incarceration", which will be done by collaborating with mental health service providers and mental health advocacy programs to promote awareness and decrease the population of diagnoses in prison (DOJ, 2009, p.1). While there are many programs such as these that serve as awareness programs for mental illness in prison, there are no currently established programs that address the issue of deepfake-related cybercrime in the metaverse.

Methods and Sample

The goal of this study is to bridge the gap from the lack of research on mitigating deepfake-related interpersonal crime in the metaverse by developing a theoretical framework for effective methods of prevention, as well as for serving support to victims who suffer from the harmful psychological impacts of these crimes. Using purposive and snowball sampling techniques, semi-structured interviews were conducted with a diverse group of metaverse experts from various sectors in South Korea (n=8), including government, private companies, and academia. Those who indicated availability were asked a series of ten interview questions relating to deepfake-related interpersonal cybercrime in the metaverse, and their responses were then used to determine an efficient framework for possible solution strategies for mitigation techniques. The average length of the interviews was 35 minutes, conducted using Zoom. However, the interviews were not recorded, as the interviewers did not receive informed consent from all the participants. To ensure an accurate analysis of responses, the interviewers took comprehensive notes during each session and followed up with the participants through email for any additional responses and or clarifications. This project was exempt from obtaining Institutional Review Board (IRB) approval due to the nature of the interviews to fit the goal of the study, which was to gather expert insights on specific policies, technologies, and techniques through the professional knowledge of participants. The interviews were coded based on the notes and email responses. To achieve intercoder reliability, a collaborative approach was used to combine the interview notes and email responses for further interpretation by the coding team. Discrepancies in interpretation were discussed and resolved, ensuring a consistent application of the coding scheme, and further mitigating potential bias. Therefore, the reliability and validity of the data analysis were consistent.

The method of including expert testimonies from South Korea derives from their advanced expertise in research and practice of the metaverse. This is supported by an issue in the New York Times (2023), which states that the government of South Korea has been actively investing in the metaverse through an innovative approach to shape the future of the platform, positioning the country as a global leader in this emerging field (Young & Stevens, 2023, para.18).

The responses of all experts were closely identified and compared to serve as a foundation for creating an effective framework for mitigating these types of crimes in the upcoming metaverse, as well as providing victim services for users who have already become victims. By collaborating these responses with prior research of traditional practices, it was determined which types of services need to be expanded to make the metaverse a safer place for the future of all users, including children, who can become victims of deepfake-related interpersonal cybercrime.

Expert Interviews

The current study includes qualitative interviews with metaverse experts (n = 8) from various professions in South Korea that were conducted in February 2023. These expert professions include the Korean Metaverse Service, the Telecommunication Service, the Police Agency of the Cyber Investigation Bureau, the National Cyber Security Center, the Software Policy Research Institute, a Professor of Cybersecurity, the Ministry of Cyber Defense, and the Ministry of Science and ICT. To ensure anonymity after transcribing their responses, the experts were assigned numbers (n= 1-8) for further analysis. The applied interview analyzing method used in this study is thematic analysis, which is a method to systematically identify, organize, and create patterns of related categories in a set of qualitative data, allowing the researcher to create a collective conclusion or finding on the data (Braun & Clarke, 2012, p.57). This approach allowed for identifying a pattern in the expert testimony on the topics of deepfake crime in the metaverse to aid in the goal of this study.

The interviewed metaverse experts emphasize the need for a theoretical framework through the application of the routine activities theory. As the metaverse is expected to evolve rapidly, it is expected to witness an increase in users of the younger generation, industrial sectors, governments, and traditional institutions within the duration of this development. A series of 8 questions were asked of the experts:

1. Regarding technological advancements of the metaverse, how do you expect to see the shift of internet users from the surface web to the metaverse in the near future?
2. What is your prediction on primary users using the metaverse in particular?
3. In your professional opinion, how do you think the potential increase of users in the metaverse will open the doors to an increase in cybercrime victimization? Who do you think their primary target is? Why do you think so?
4. Specifically, interpersonal cybercrime victimization (ID theft, cyber-sexual crime, cyberstalking/cyber harassment): How significant is this issue in comparison to what it will become? What is your opinion on children becoming targets of these crimes?
5. Who is the typical cybercriminal in the metaverse? What do you think their motives are?
6. Do you think that there is a reason for their motive to commit interpersonal cybercrime, on this platform, perhaps, a lack of capable guardianship? If so, how do you suggest action to be made for an increase in capable guardianship?
7. What are your thoughts on the rise of deepfake technology used to commit interpersonal cybercrime (as seen on the surface web) migrating to the metaverse? What measures (legally, policy, technologically) should be taken to prevent deepfake from taking over the metaverse?
8. Interpersonal cybercrime in the metaverse and on the surface web can intensely impact the victim's psychological health. In response, what are your thoughts on whether a potential solution will include mental health services provided within the virtual world to assist the victims of these crimes committed in the metaverse?

Results

Motivated Offender

In response to the Q5 on the typical cybercriminal in the metaverse, experts commented on the age, motives, anonymity, guilt, lack of restrictions, and specialization of these cybercriminals.

Expert 1 suggests that the age of these motivated offenders is likely to be in their 20s. It was further suggested that the crimes of this age group with strong sexual or economic needs with low social achievement will increase in the metaverse, as it is an entirely online platform. In addition, there are concerns about increasing crime rates as an increase in opportunity for motivated offenders to commit crimes will result as the metaverse develops. This is because users are expected to form close communication and curiosity with each other on the metaverse platform.

3 of the experts agreed that the motivation for these crimes will include financial gain or sexual gratification. Specifically, cyber offenders who seek deepfake related cyber-sexual interpersonal crimes are motivated by their need to take sexual basking, as suggested by Expert 2, who claims that *"besides financial gain or sexual motive, there is no particular reason to commit a crime between users in the metaverse"*. An additional motive that was established through these interviews includes the lack of restrictions in the metaverse"

An additional motive that was established through these interviews includes the lack of restrictions in the metaverse, as more offenders will be likely to act upon their supportive motives to commit interpersonal cybercrime in virtual reality.

Expert 7 associates this motivation with cryptocurrency hacking in North Korea, testifying that cybercriminals will consider the metaverse as their playground for financial crime, as it is vulnerable and lacks security measures and legislation. It was further suggested that monetary gain is the primary motivation of these cyber offenders who commit deepfake financial interpersonal cybercrime in the metaverse. Expert 6 suggests that money extortion along with other financial crimes is expected to become prevalent in the upcoming metaverse. Furthermore, it was suggested that the motive of monetary gain is carried out through the cybercrime of money extortion as a result of the "Nth room" (Expert 6), a South Korean cybercrime scandal involving sextortion and blackmailing of women in cyberspace via the application Telegram (Joohee & Chang, 2021, p.1).

Motivated offenders ensure their anonymity using avatars and deepfake-manipulated users, which are the result of deepfake technology to clone or impersonate another avatar (Tariq et al., 2023). Expert 3 expresses the need for preventative measures to be established, as *"it can be problematic if a metaverse user is tempted to deviation behind the anonymity that one doesn't show their natural face"*. Therefore, the anonymity provided through using avatars and deepfake manipulated users in the metaverse can encourage users to commit these interpersonal crimes. Anonymity also decreases the guilt of cyber-offenders, as they also do not associate their identity with their avatar.

As traditional profiling for cybercriminals on the surface web provided a framework for a specific profile of motivated offenders, the same cannot be done for metaverse criminals. However, we can make correlations using the specializations of similar traditional cybercrime. For example, Expert 8 also testified that the motivation for specialization is similar amongst traditional cybercrime and those in the metaverse, as altered to be suitable with characteristics of the metaverse such as new socialization and values.

Suitable Target

Q3 inquired about whom the primary target of cybercrime is in the metaverse. Victims of cybercrime in the metaverse are expected to include the younger generation (specifically teenagers) (Expert 1, Expert 2, Expert 3, Expert 5) aged between 10 and 20 years old (Expert 4, Expert 7) the elderly (Expert 2), and women (Expert 4). Previous research has found that approximately 61% of individuals ages 10-18 experience cyber victimization (Garthe et al., 2023, p.1). Specific to deepfake, women are primary targets for motivations of interpersonal sexual crime, although men experience victimization as well (Rousay, 20203, p.96). The metaverse is expected to also have a high rate of interpersonal crimes such as sexual crime, invasion of privacy, and personal data infringement. Expert 3 stresses the case of deepfake-related crime in the metaverse to target the younger population, as the avatar's gender, voice, and appearance can be camouflaged using deepfake technology. It is also emphasized by Expert 8, that the main threat of deepfake-related interpersonal cybercrime in the metaverse is identity-related, as one or multiple avatars can manipulate their appearance by cloning another user and therefore steal assets such as cryptocurrency and/or NFTs. Expert 8 further explains that *"if someone can clone or pretend to be another user's avatar, they can become the new leader of that group or the new friends of friends. This means that identity theft or personification of someone else will be the primary target"*.

When asked about the suitable targets of interpersonal cybercrime victimization in the metaverse (Q4), experts held a consensus about children (Expert 2, Expert 3, Expert 4, Expert 5, Expert 7) and teenagers (Expert 1 and Expert 7) being primary targets in the deepfake facilitated interpersonal crimes such as cyber grooming, cyber stalking, and cyber harassment. It was expressed that since child sexual cybercrimes are prevalent on the surface web, it can be expected to see similar criminal behavior in the metaverse (Expert 2, Expert 3, Expert 4, Expert 5, and Expert 7). Specifically, children who are unsupervised or have weaker judgments will likely fall victim to cybercrime. Expert 4 suggests that a systematic approach is necessary to mitigate child harassment crimes in the metaverse, with the goal to detect problems early and impose restrictions or sanctions for this behavior. Most of the remaining expert's responses (Expert 2, Expert 4, Expert 5, Expert 7, and Expert 8) suggest that interpersonal cybercrime victimization in the metaverse is a significant issue, as these crimes on this platform will essentially imitate those on the surface web in a more complex way (Expert 8). However, Expert 1 expressed a conflicting perspective, suggesting that crime on the metaverse will differ from the surface web. Nonetheless, it was agreed that these crimes require immediate attention and action from policymakers and stakeholders to prevent their proliferation in the metaverse. In summary, the need for a theoretical framework is urgent and should be implemented before interpersonal cybercrime takes over the metaverse.

Capable Guardianship

Through expert testimony, it was found that there was a consensus regarding the use of proper guardianship as an effective way to mitigate cybercrime in the metaverse. However, it was argued that since crime seems to occur on the surface web even in the presence of a capable guardian, there is a chance that guardians do not greatly assist in preventing and reducing crime (Expert 6). Despite this argument, it was agreed that the presence of guardians can be deemed effective in mitigating cybercrime in various ways, through informal and formal guardianship measures.

Informal Capable Guardian

Methods of an informal capable guardian suggested by metaverse experts include implementing measures to protect user surveillance and privacy, establishing metaverse ethics and user protection policies, proper training of higher authority in the metaverse (schools, parents, teachers), and establishing a customer service desk that is available for contact 24/7. A balanced approach to legal and institutional arrangements was also suggested, in addition to the introduction of a Real ID authentication system to ensure the user is controlling only their avatar without the capability of using deepfake technology to manipulate their own user or anyone else's.

Formal Capable Guardian

The interviewees shared their expertise in cybersecurity by providing suggestions on potential solutions for achieving the goal of mitigating cybercrime in the upcoming metaverse as well as providing support for victims' psychological health. It was found that formally capable guardians such as legal and policy methods could be very useful for acting against cybercrime in virtual reality.

It was found through interview responses that in correlation to the current laws on cybercrime, criminal laws have been established for offenses that cause economic damage to the victim, rather than emotional harm. On the other hand, policy measures have not been addressed in response to psychological distress and emotional harm in cybercrime victims. Therefore, it is extremely important to recognize the psychological health of cybercrime victims, especially in the metaverse, and take action by providing a framework of support alternatives.

Experts also stressed the importance of prioritizing victims' psychological health and suggested that legal action is imperative to be taken for all cybercrime offenses in the metaverse. Expert 2 discussed that legal assistance should be provided for all crimes, as well as the opportunity for victims to seek counseling and psychiatric therapy for sexual offenses, as they cause extreme emotional distress. Experts 7 and 8 agreed that psychological and public health considerations are essential in establishing preventative actions, as well as supporting victims mental recovery through healing programs.

Policy Implications

Policy and Legal Framework

Analysis of expert testimony suggests many ways to support victims of interpersonal cybercrime in the metaverse through policy and legal assistance. Prior to addressing the actions taken for metaverse cybercrime, a clear definition of the metaverse account and owner should be established with an implementation of a regular monitoring system that always displays the user's identity (Expert 4). This should be done with the intention of detecting cybercriminal behavior by complying with the current legal system (Expert 6). Furthermore, users should be thoroughly educated in cyber-ethics (Expert 4). This ensures that prior to the user interacting with others in the metaverse, they are completely aware of the consequences of criminal activity and the expectations of acceptable behavior in the metaverse.

The first step towards implementing a policy and legal framework for criminal laws of cybercrime should be established, as current cybercrime laws focus mainly on the impacts of economic harm, rather than emotional harm (Expert 1). Experts stress their concern for the high risk of deepfake technology being used in the metaverse, as it is already occurring on the surface web (Expert 1, Expert 2, Expert 3, Expert 4, Expert 5, Expert 6). Therefore, legal action is suggested to be taken to the same standards in the real world when these established laws are broken by physically arresting the cybercriminal and following court procedures. Once these laws are established and the cybercriminal is held accountable, legal assistance, financial aid, counseling, and psychiatric therapy for emotional distress can be provided to the victim.

Cybercrime victim advocacy and assistance programs are to be provided in the metaverse in accordance with those which are already in operation in the real world to ensure proper support and healing. In addition to including these victim assistance programs, it is suggested that the public proactively run healing programs to rehabilitate the victims' mental health, with the support and cooperation of the private sector (Expert 7). The metaverse should also implement similar victim support programs to assist victims with personal rehabilitation needs (Expert 2). Such victim support programs that are successful in the physical world should be implemented in the metaverse to provide support and healing (Expert 6). The ramification of cybercrime can affect the lives of individuals in the physical world, as cybercrime in the metaverse is not separated from the real world (Expert 7). Therefore, victims would benefit greatly from receiving this support through the metaverse to ensure their future protection and comfortability using the platform. Since there are limited specialized institutions to assist cybercrime victims in the Metaverse (Expert 3), developing a metaverse that helps victims of cybercrime is possible (Expert 5), and important.

Public Awareness Framework

There is little current established public awareness on the issue of deepfake creation in the metaverse. As the metaverse is becoming increasingly popular, it is important to implement awareness techniques to further prevent the use of illegal activity on this platform. This study suggests creating a safe space on the platform for users to socialize using technological and traditional methods to promote responsible behavior in the metaverse through awareness-raising initiatives (Expert 8).

Technological Framework

The possibility of cloning an avatar and creating a deepfake is becoming easier for users of the metaverse. Metaverse experts describe biometrics as a sufficient technological authentication technique, however, it will not prevent the creation of avatar clones and deepfake in the metaverse.

Metaverse experts firmly believe that to prevent deepfake crimes, it is essential to have access to technology that can detect manipulated faces. Since there is not yet a method to detect deepfakes in the metaverse, related studies provide a guideline for us to establish a theoretical framework for deepfake detection in virtual reality. Deepfakes using Generative Adversarial Networks (GAN) are the most common form of deepfake, which are created through the neural networks: 'Generator', which replicates data sampling, and 'Discriminator', which detects the data as being original (real), or fake (Guarnera, Giudice & Battiato, 2023, p.1). Additionally, a previous study by Guarnera et al., (2023) found that using a single ResNET-34 encoder to detect images using a hierarchal approach of 3 levels will lead to successful results in detecting and explaining the details of an AI-generated image (deepfake) by eliminating the real data and revealing the fake data (p.3).

Previous studies found a new deepfake detection system, SOTA (state-of-the-art) techniques to be successful in deepfake detection. One challenge of this technique is how to identify the model of deepfake creation. It was therefore proposed by Guarnera et al., (2022), to develop a model recognition, which will trace the artificial image to the generator, or model owner. This can be particularly helpful in the metaverse, as the deepfake image can be traced to the metaverse user who created it (Guarnera et al., 2022, p.1).

Discussion

The current study provides a suggestion for a developed framework for professionals to use in mitigating deepfake interpersonal crime in the metaverse and further provides support for victims who suffer from psychological damage as a result. As the metaverse continues to develop, the need for mitigation and intervention of cybercrime will increase, with deepfake existing as a major threat. For this reason, this study focuses on the risks of deepfake in the metaverse and provides a framework for intervention and prevention.

8 metaverse experts from South Korea were chosen through a snowball sampling method to participate in this study. The primary focus on South Korean experts is attributed to the nation's active government support for the metaverse projects across a range of industries, as well as the recent development of numerous metaverse platforms. These expert testimonies served as a foundation for this study's management framework, as capable guardianship solutions of legal, policy, awareness, and technological methods were established along with the support of previous literature. It was found that deepfake is a rising threat to the

users of the metaverse, specifically children and teenagers, which calls for a need to establish criminal procedures, police enforcement, and further provide psychological healing programs to victims impacted by these crimes.

These findings are important for criminology professionals in all arenas: academia, crime prevention, policymakers, and the metaverse's own user policy network. This study serves as the foundation on how to address the issue of deepfake interpersonal crime in the metaverse and is open to development as the metaverse continues to advance. The combined application of Eysenck's theory of criminality and the routine activities theory helps define a motivated offender through psychological wellness. The findings of this study suggest that the routine activities theory can also be successfully applied to creating a management framework, as each component served a great purpose to this study: (1) understanding a motivated offender, (2) understanding a suitable target, and further victimization risks, and especially (3) capable guardianship, which served as a guideline for this study's framework architecture. The use of RAT's capable guardianship assisted in developing appropriate formal and informal control mechanisms for managing deepfake crime in the metaverse such as the implementation of criminal laws and prosecution, law enforcement security, and community guidelines.

Limitations of this study include the lack of previous knowledge determining the success of deep fake mitigation in the metaverse to serve as a foundation for developing successful measures. An additional limitation of this study is the limited perspective of expert testimony. Since this study only interviewed experts from South Korea, this research is very specific and may not result in the same report if testimonies were included from additional countries.

However, the lack of previous literature on this topic functions as a strength, as this study fills the gap of the lack of research between deepfake-facilitated interpersonal crime in the metaverse and approaches to the harmful psychological effects of these crimes. This study is important as it is the first study to acknowledge the issue of deepfake-related interpersonal crime in the metaverse from a psychological perspective. Additionally, South Korean expert testimony is an extreme strength of this research as South Korea is known to be the global leader in technology and digital governance (Chung, 2020, p.1). As we expect to see a rise in similar practices in the United States, these experts' testimonies served as a strong foundation for a developmental framework of mitigation and prevention.

References

Ahmed, B., Ali, G., Hussain, A., Baseer, A., & Ahmed, J. (2021). Analysis of text feature extractors using deep learning on fake news. *Engineering, Technology & Applied Science Research, 11*(2), 7001-7005

Anderson, J., & Rainie, L. (2022). The metaverse in 2040. *Pew Research Centre.*

Bale, A. S., Ghorpade, N., Hashim, M. F., Vaishnav, J., & Almaspoor, Z. (2022). A Comprehensive Study on Metaverse and Its Impacts on Humans. *Advances in Human-Computer Interaction*, 2022.

Braun, V., & Clarke, V. (2012). Thematic analysis. *American Psychological Association.*

Broll, R., & Huey, L. (2015). "Just being mean to somebody isn'ta police matter": Police perspectives on policing cyberbullying. *Journal of school violence, 14*(2), 155-176.

Bui, J. (2022). The metaverse is one giant deep fake. *LinkedIn.* Retrieved March 7, 2023, from https://www.linkedin.com/pulse/metaverse-one-giant-deep-fake-jerry-bui-cfe/

Castillo, M. del. (2022, September 1). Facebook's metaverse could be overrun by deep fakes and other misinformation if these non-profits don't succeed. *Forbes*. Retrieved March 7, 2023, from https://www.forbes.com/sites/michaeldelcastillo/2022/08/29/facebooks-metaverse-could-be-overrun-by-deep-fakes-and-other-misinformation-if-these-non-profits-dont-succeed/\

Cohen, L. E., & Felson, M. (1979). Social change and crime rate trends: A routine activity approach. *American sociological review*, 588-608.

Coventry, L., Briggs, P., Jeske, D., & Van Moorsel, A. (2014). SCENE: A structured means for creating and evaluating behavioral nudges in a cyber security environment. In Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience: Third International Conference, DUXU 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part I 3 (pp. 229-239). *Springer International Publishing.*

Cheong, B. C. (2022). Avatars in the metaverse: potential legal issues and remedies. *International Cybersecurity Law Review*, 1-28.

Choi, K. S. (2008). Computer crime victimization and integrated theory: An empirical assessment. *International Journal of Cyber Criminology, 2*(1).

Choi, K. S., Cho, S., & Lee, J. R. (2019). Impacts of online risky behaviors and cybersecurity management o-n cyberbullying and traditional bullying victimization among Korean youth: Application of cyber-routine activities theory with latent class analysis. *Computers in Human Behavior*, 100, 1-10.

Chung, C. S. (2020). Developing digital governance: South Korea as a global digital government leader. *Routledge.*

Das, A. (2022, February 8). Woman recalls 'gang rape' in metaverse; concerns grow over making VR platfor-ms safe from sexual predators. cnbctv18.com. Retrieved March 7, 2023, from https://www.cnbctv18.com/technology/woman-recalls-gang-rape-in-metaverse-concerns-grow-over-making-vr-platforms-safe-from-sexual-predators-12396992.htm

Dick, E. (2021, November 15). Public policy for the metaverse: Key takeaways from the 2021 AR/VR Policy Conference. Public Policy for the Metaverse: Key Takeaways from the 2021 AR/VR Policy Conference. Retrieved March 7, 2023, from http://www.iitfm.org/public-policy-metaverse-key-takeaways-2021-arvr-policy-conference.html

DOJ. (2017, April 7). Addressing mental illness in the criminal justice system. *The United States Department of Justice*. Retrieved March 7, 2023, from https://www.justice.gov/archives/opa/blog/addressing-mental-illness-criminal-justice-system

Eysenck, H. J., & Gudjonsson, G. H. (1989). The causes and cures of criminality. S*pringer Science & Business Media.*

Ferraro, M. F. (2019). Deepfake legislation: A nationwide survey-state and federal lawmakers consider legislation to regulate manipulated media. online], WilmerHale https://www. idsupra. com/legalnews/deepfake-legislation-anationwide-86809.

Galer, S. (2022, October 3). SAP brandvoice: Why the metaverse is a Gutenberg moment for business. *Forbes*. https://www.forbes.com/sites/sap/2022/09/27/why-the-metaverse-is-a-gutenberg-moment-for-business/?sh=34b3ae663259

Gamiz, M. (2021). News post: Chalfont woman ordered to face trial for cyber harassing three teens on. *CIMEWATCH*. Retrieved March 7, 2023, from https://bucks.crimewatchpa.com/da/29567/post/chalfont-woman-ordered-face-trial-cyber-harassing-three-teens-daughter%E2%80%99s-cheer-squad

Garthe, R. C., Kim, S., Welsh, M., Wegmann, K., & Klingenberg, J. (2023). Cyber-victimization and mental health concerns among middle school students before and during the COVID-19 pandemic. *Journal of youth and adolescence, 52*(4), 840-851.

Goldsmith, A. J. (1981). Eysenck's Theory of Criminal Personality-A Review of Recent Evidence and the Implications for Criminological Theory and Social Practice. *In Canadian Criminology Forum* (Vol. 4, No. 2, pp. 88-102).

Guarnera, L., Giudice, O., & Battiato, S. (2023). Level Up the Deepfake Detection: a Method to Effectively Discriminate Images Generated by GAN Architectures and Diffusion Models. arXiv preprint arXiv:2303.00608.

Guarnera, L., Giudice, O., Nießner, M., & Battiato, S. (2022). On the Exploitation of Deepfake Model Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 61-70).

Gudjonsson, G. (2016). Hans Eysenck's theory on the 'causes' and 'cures' of criminality: A personal reflection. *Personality and Individual Differences, 103*, 105-112.

Guo, H., Wang, X., & Lyu, S. (2022). Detection of Real-time DeepFakes in Video Conferencing with Active Probing and Corneal Reflection. arXiv preprint arXiv:2210.14153.

Han, J. (2022). An information ethics framework based on ICT platforms. *Information, 13*(9), 440.

Harrell, D. F. (2022, June 25). Beyond the 'metaverse': Empowerment in a blended reality. Open Learning. Retrieved March 7, 2023, from https://openlearning.mit.edu/news/beyond-metaverse-empowerment-blended-reality

Joohee, K., & Chang, J. (2021). Nth room incident in the age of popular feminism: a big data analysis. *Azalea: Journal of Korean Literature & Culture, 14*(14), 261-287.

Kasiyanto, S., & Kilinc, M. R. (2022). The Legal Conundrums of the Metaverse. *Journal of Central Banking Law and Institutions, 1*(2), 299-322.

Katie, K. (2022, June 9). Bucks county mother gets probation in harassment case involving daughter's cheerleading rivals. *6abc Philadelphia*. Retrieved March 7, 2023, from https://6abc.com/raffaela-spone-bucks-county-pa-cheerleaders-harassment-case-victory-vipers-squad/11939419/

Kirwan, G., & Power, A. (2013). Cybercrime: The psychology of online offenders. *Cambridge University Press.*

Kürtünlüoğlu, P., Akdik, B., & Karaarslan, E. (2022). Security of virtual reality authentication methods in metaverse: An overview. arXiv preprint arXiv:2209.06447.

Leukfeldt, R., & Holt, T. J. (Eds.). (2019). The human factor of cybercrime. *Routledge.*

Liao, S. (2022, September 9). Roblox wants to advertise to gamers ages 13 and up in the metaverse. *The Washington Post*. Retrieved March 7, 2023, from
https://www.washingtonpost.com/video-games/2022/09/09/roblox-ads-metaverse/

Lickiewicz, J. (2011). Cyber Crime psychology-proposal of an offender psychological profile. *Problems of forensic sciences, 2*(3), 239-252.

Meskys, E., Kalpokiene, J., Jurcys, P., & Liaudanskas, A. (2020). Regulating deep fakes: legal and ethical considerations. *Journal of Intellectual Property Law & Practice, 15*(1), 24-31.

Moioli, F. (2022, October 12). Council post: The metaverse: Don't confuse it with virtual reality. *Forbes.* https://www.forbes.com/sites/forbestechcouncil/2022/08/11/the-metaverse-dont-confuse-it-with-virtual-reality/?sh=3e6e393b2257

Montasari, R., Hosseinian-Far, A., & Hill, R. (2018). Policies, innovative self-adaptive techniques and understanding psychology of cybersecurity to counter adversarial attacks in network and cyber environments. *Cyber criminology*, 71-93.

Mohsin, K. (2021). The Internet and its Opportunities for Cybercrime–Interpersonal Cybercrime. *SSRN Electronic Journal*, 2021. URL: https://doi. org/10.2139/ssrn. 3815973 (date of request: 28.02. 2022).

Parti, K. (2011). Actual Policing in Virtual Reality-A Cause of Moral Panic or a Justified Need?. In Virtual Reality. *InTech Open.*

Pepper, C. (2022, October 7). Lawyer: What to do about Deepfake and metaverse libel threat to publishers. *Press Gazette*. Retrieved March 7, 2023, from https://pressgazette.co.uk/comment-analysis/deepfake-metaverse-publishers/

Peterson, J. K., Skeem, J., Kennealy, P., Bray, B., & Zvonkovic, A. (2014). How often and how consistently do symptoms directly precede criminal behavior among offenders with mental illness?. *Law and human behavior, 38*(5), 439.

Pettifer, S., Barrett, E., Marsh, J., Hill, K., Turner, P., & Flynn, S. (2022, January 1). The future of Extended Reality Technologies, and implications for online child sexual exploitation and abuse. *Semantic Scholar*. Retrieved March 7, 2023, from https://www.semanticscholar.org/paper/The-Future-of-eXtended-Reality-Technologies%2C-and/d923e71c4f932398faf86516db0eca635e509be7

Rousay, Victoria. 2023. Sexual Deepfakes and Image-Based Sexual Abuse: Victim-Survivor Experiences and Embodied Harms. *Master's thesis, Harvard University Division of Continuing Education*.

Saklofske, D. H., McKerracher, D. W., & Eysenck, S. B. (1978). Eysenck's theory of criminality: a scale of criminal propensity as a measure of antisocial behavior. *Psychological Reports, 43*(3), 683-686.

Stalans, L. J., & Donner, C. M. (2018). Explaining why cybercrime occurs: Criminological and psychological theories. *Cyber Criminology*, 25-45.

Suzor, N. (2010). The role of the rule of law in virtual communities. *Berkeley Tech. LJ, 25*, 1817.

Tariq, S., Abuadbba, A., & Moore, K. (2023). Deepfake in the Metaverse: Security Implications for Virtual Gaming, Meetings, and Offices. arXiv preprint arXiv:2303.14612.

Usmani SS, Sharath M, Mehendale M. Future of mental health in the metaverse. General Psychiatry 2022; 35:e100825. doi:10.1136/gpsych-2022-100825

Whittaker, L., Kietzmann, T. C., Kietzmann, J., & Dabirian, A. (2020). "All around me are synthetic faces": the mad world of AI-generated media. *IT Professional, 22*(5), 90-99.

Woods, N. (2022). Users' Psychopathologies: Impact on Cybercrime Vulnerabilities and Cybersecurity Behavior. In Cyber Security: Critical Infrastructure Protection (pp. 93-134). *Cham: Springer International Publishing*.

Yang, R., Li, L., Gan, W., Chen, Z., & Qi, Z. (2023, April). The Human-Centric Metaverse: A Survey. In Companion Proceedings of the ACM Web Conference 2023 (pp. 1296-1306).

Young, J. Y., & Stevens, M. (2023, January 29). Will the metaverse be entertaining? ask South Korea. *The New York Times*. Retrieved April 29, 2023, from https://www.nytimes.com/2023/01/29/business/metaverse-k-pop-south-korea.html