A Comprehensive Review of Security and Privacy Challenges and Solutions in Autonomous Driving Systems

Mohammed Amin [1], **Youakim Badr[2]**, **Qais Al-Na'amneh[3]**, **Mahmoud Aljawarneh[3]**, **Rahaf Hazaymih[4]**,
Shahid Munir Shah[5]

[1] *Fellowship Researcher, INTI International University, Nilai 71800, Malaysia*
[2] *School of Graduate Professional Studies, The Pennsylvania State University, Malvern, PA, USA*
[3] *Faculty of Information Technology, Applied Science Private University, Amman, Jordan*
[4] *Dept. Computer Science, Jordan University of Science and Technology, Irbid, Jordan*
[5] *Faculty of Engineering Sciences and Technology, Hamdard University, Karachi, Pakistan*

ARTICLE INFO

ABSTRACT

The rapid evolution of immersive technologies such as Augmented Reality (AR) and Virtual Reality (VR) has transformed sectors ranging from entertainment and healthcare to education and industrial operations. However, the increasing integration of these technologies into daily life introduces a new landscape of cybersecurity and privacy challenges. This review paper provides a comprehensive comparative analysis of security threats associated with AR and VR systems, emphasizing the unique vulnerabilities that arise from their distinct architectures and user interaction models. We examine emerging attack vectors such as sensor spoofing, man-in-the-room attacks, data leakage through AR overlays, VR hijacking, and unauthorized motion tracking. The paper also explores crosscutting issues like biometric data misuse, identity theft in virtual spaces, and spatial-temporal data inference. A critical comparison is made between the threat surfaces of AR—where the virtual is overlaid on the physical—and VR—where users are fully immersed in synthetic environments. In parallel, we evaluate a range of mitigation strategies and defense mechanisms, including secure sensor integration, encryption protocols, context-aware access control, and privacy-preserving rendering techniques. The paper concludes by identifying key research gaps and proposing a roadmap for developing holistic and resilient security frameworks tailored to the future of immersive technologies.

Keywords: Augmented Reality, Virtual Reality, Cybersecurity, Privacy, Sensor Spoofing, Data Leakage, VR Hijacking, Motion Tracking, Biometrics, Mitigation Strategies, Immersive Technologies.

How to cite the article

1. Introduction

Immersive technologies, encompassing Augmented Reality (AR) and Virtual Reality (VR), represent a paradigm shift in human-computer interaction, promising to redefine how individuals perceive, interact with, and manipulate digital information interwoven with their physical surroundings or within entirely synthetic worlds. AR overlays digital content onto the user's view of the real world, enhancing physical reality with contextual information, interactive elements, or virtual objects. Conversely, VR completely immerses the user in a computer-generated environment, replacing the physical world with a synthetic one, often facilitated through head-mounted displays (HMDs) and sensory feedback mechanisms. The potential applications are vast and transformative, spanning entertainment, healthcare simulations and interventions, remote collaboration, industrial training and maintenance, automotive design [1], and educational experiences.

This accelerating integration into diverse facets of life, however, is accompanied by a burgeoning array of security and privacy concerns that are fundamentally distinct from traditional computing paradigms. The intimate connection between the user's physical actions, sensory inputs, and the digital realm within AR/VR systems creates novel attack surfaces and amplifies the potential impact of security breaches as shown in Figure 1. [2]. Unlike conventional systems where attacks might compromise data or disrupt services, attacks on immersive systems can directly manipulate a user's perception of reality, steal highly sensitive biometric and behavioral data, or even induce physical discomfort or harm [3].
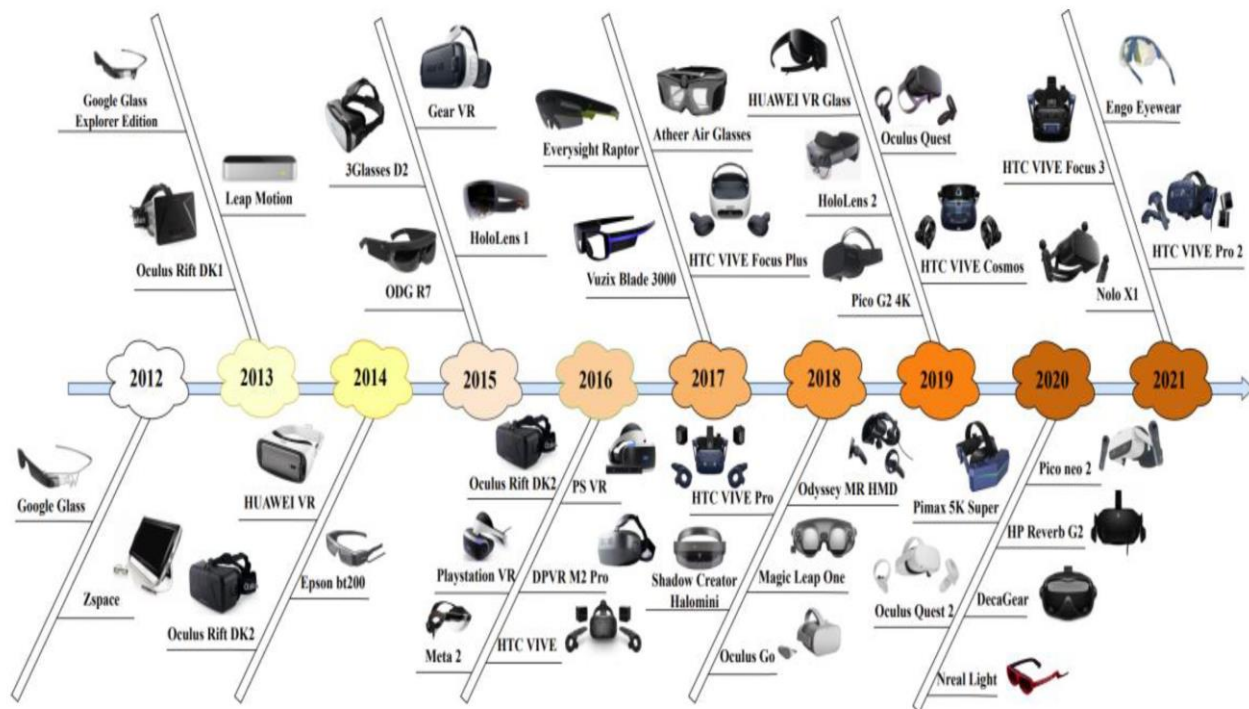

Figure 1. Overview Augmented and Virtual Reality

The unique characteristics of AR and VR necessitate a specialized examination of their respective security vulnerabilities. AR systems, by their nature, bridge the physical and digital, making the integrity of sensor data mapping the real world (cameras, LiDAR, GPS, and IMUs) paramount. Attacks targeting these sensors can corrupt the user's perception of their immediate physical environment, leading to potentially hazardous situations. Furthermore, the overlay mechanism itself presents risks of data leakage or malicious information injection [4]. VR systems, while isolating users from the physical world, create vulnerabilities related to the complete control an attacker might gain over the user's perceived reality. Hijacking the virtual environment, manipulating sensory feedback, or exploiting the rich motion and biometric data collected for interaction can lead to profound psychological manipulation, identity theft, or unauthorized surveillance.

This paper presents a comprehensive comparative analysis of the security and privacy threats inherent in AR and VR technologies. We delve into the specific attack vectors facilitated by their distinct architectures and user interaction modalities, moving beyond generalized cybersecurity principles to address the nuances of immersive systems. Key contributions include:

• A detailed examination of AR and VR system architectures, highlighting components and data flows pertinent to security analysis.

• A critical comparison of the threat surfaces unique to AR (physical-digital interface) and VR (synthetic immersion).

• In-depth exploration of emerging attack vectors, including sensor data manipulation, reality distortion attacks, and sophisticated tracking techniques.

• Analysis of cross-cutting challenges such as the exploitation of biometric identifiers intrinsically captured by these systems and the potential for identity fraud within virtual environments.

• Evaluation of current and proposed mitigation strategies, assessing their applicability and limitations in the context of AR/VR.

• Identification of critical research gaps and the proposal of a forward-looking roadmap towards establishing robust, resilient security frameworks for the future of immersive computing.

By synthesizing current knowledge and highlighting unresolved challenges, this work aims to inform researchers, developers, policymakers, and users about the critical need for proactive security and privacy measures as AR and VR technologies become increasingly ubiquitous. The development of trustworthy immersive experiences hinges upon our collective ability to anticipate, understand, and neutralize these evolving threats.

2. Background: AR and VR Architectures and Interaction Models

Understanding the security implications of AR and VR requires a foundational grasp of their underlying architectures, core components, and modes of user interaction. While often discussed together under the umbrella of Extended Reality (XR) [5], AR and VR possess distinct technological underpinnings that shape their respective vulnerabilities. Table I summarizes the research work in AR and VR related studies.

2.1 Defining AR and VR

Augmented Reality enhances, rather than replaces, the user's perception of the physical world. It achieves this by overlaying computer-generated sensory information (visual, auditory, haptic) onto the real-world environment. Key characteristics typically include [6]: (1) Combination of real and virtual worlds. (2) Real-time interaction. (3) Registration of virtual objects within the 3D real-world context. AR systems range from handheld mobile devices using camera feeds (e.g., Pokemon GO, IKEA Place) ´ to sophisticated head-mounted displays (HMDs) like Microsoft HoloLens or Magic Leap, which employ see-through optics or video pass-through mechanisms. Virtual Reality aims to create a sense of presence within a completely synthetic, computer-generated environment. It isolates the user from the physical world, primarily through opaque HMDs (e.g., Oculus Rift/Quest, HTC Vive, and PlayStation VR) that replace the user's visual field with rendered scenes [8]. Auditory immersion is achieved via headphones delivering spatialized sound, and interaction often involves handheld controllers tracked in 3D space, or increasingly, hand and body tracking. The goal is to achieve a high degree of immersion and plausibility within the virtual simulation [10].

2.2 Key Architectural Components

Both AR and VR systems share common functional blocks as shown in Fig 2, but their emphasis and implementation differ significantly, impacting the security landscape, Table II provides the summary of Key Architectural Components. A generalized architecture includes: Sensors: The primary interface for capturing information about the user and their environment.

– Common: Inertial Measurement Units (IMUs) for orientation tracking (accelerometers, gyroscopes, magnetometers) [11].

– AR-centric: Outward-facing cameras (RGB, depth), LiDAR scanners for environmental mapping and Simultaneous Localization and Mapping (SLAM) [12], GPS/GNSS for location, microphones for ambient sound and voice commands.

– VR-centric: Inside-out or outside-in tracking cameras (infrared or visible light) for positional tracking of the HMD and controllers [13], eye-tracking cameras within the HMD [14], microphones for voice interaction.

– Emerging: Physiological sensors (EEG, EMG, EKG, GSR) integrated into HMDs or peripherals for affective computing or health monitoring [15].

• Compute Unit: Processes sensor data, runs applications, renders virtual content, and manages system resources. This can be integrated into the HMD, tethered to a PC or console, or increasingly rely on edge or cloud computing for complex tasks (e.g., object recognition, remote rendering). The distribution of computation introduces network dependencies and associated vulnerabilities.

• Display Technology: Presents visual information to the user.

– AR: Optical see-through displays (waveguides, reflective combiners) or video pass-through displays (using external cameras to show the real world on an internal screen with overlays) [16]. Each has different latency and registration challenges affecting security [17].

– VR: Opaque displays (OLED, LCD) presenting stereoscopic images to create depth perception. High resolution and refresh rates are crucial for immersion and mitigating motion sickness [18].

• Interaction Devices: Allow user input and manipulation.

– Handheld controllers (6DoF tracked).

– Hand tracking via cameras.

– Eye tracking for gaze-based interaction or foveated rendering.

– Voice command recognition.

– Haptic feedback devices (gloves, vests).

• Software Stack: Includes the operating system (often based on Android or proprietary systems), middleware (tracking libraries, rendering engines like Unity or Unreal Engine), and Software Development Kits (SDKs), and end-user applications. Vulnerabilities can exist at any layer of this stack [19].

Table 1. The comparison of the related studies in defining AR and VR

Reference	Main Focus	Key Contribution/Findings	Limitations/Future Work (as identified)
[5]	Role of Machine Learning in XR	Explores the application and integration of machine learning techniques within Extended Reality systems.	Discusses challenges in data needs, model accuracy, and real-time processing constraints.
[7]	Security & Privacy in AR	Seminal work identifying fundamental security and privacy challenges in AR, including sensor access, output manipulation, and user perception manipulation risks.	Primarily identifies challenges; proposes areas for future research rather than definitive solutions.
[6]	XR Security & Privacy in Healthcare	Scoping review analyzing threats in XR healthcare, highlighting risks from extensive data collection (biometrics, behavior) and specific XR attack vectors (e.g., chaperone attacks). [1]	Focused on the healthcare domain. Suggests need for robust security frameworks and ethical guidelines.
[8]	Security/Privacy in Mixed Reality (MR)	Surveys security and privacy issues specifically relevant to Mixed Reality (MR), which encompasses aspects of both AR and VR. [3]	Scope generally limited to MR; practical implementations of defenses often lag behind threat identification

[9]	Security/Privacy in MultiUser AR	Explores user perceptions and identifies security/privacy risks (like unwanted recording, view interference) in scenarios where multiple users interact in a shared AR environment, using HoloLens as a case study. [4]	Findings based on specific hardware (HoloLens) and limited user groups. Focused on identifying rather than solving problems.

2.3 Data Flow and Processing

A typical data flow involves: 1. Sensing: Continuous collection of data from various sensors (user motion, environmental features, and user biometrics). 2. Processing/Tracking: Sensor fusion algorithms combine data (e.g., IMU + camera data for SLAM) to estimate the user's pose (position and orientation) and map the environment (AR) or track controllers/limbs (VR). This stage is computationally intensive and critical for system stability. 3. Application Logic: The running application determines what virtual content to display or how to respond based on user input, pose, and potentially external data sources. 4. Rendering: The compute unit generates the appropriate images (stereoscopic for VR, overlaid for AR) based on the application state and user's viewpoint. Techniques like foveated rendering optimize this process. 5. Display/Feedback: The rendered images are sent to the display, accompanied by synchronized audio and haptic feedback where applicable. This continuous loop operates at high frequencies (typically 60-140 Hz) to maintain immersion and prevent discomfort [22]. Any disruption or manipulation within this loop can severely impact the user experience and introduce security risks. Table 3 provides the summary of Data Flow and Processing.
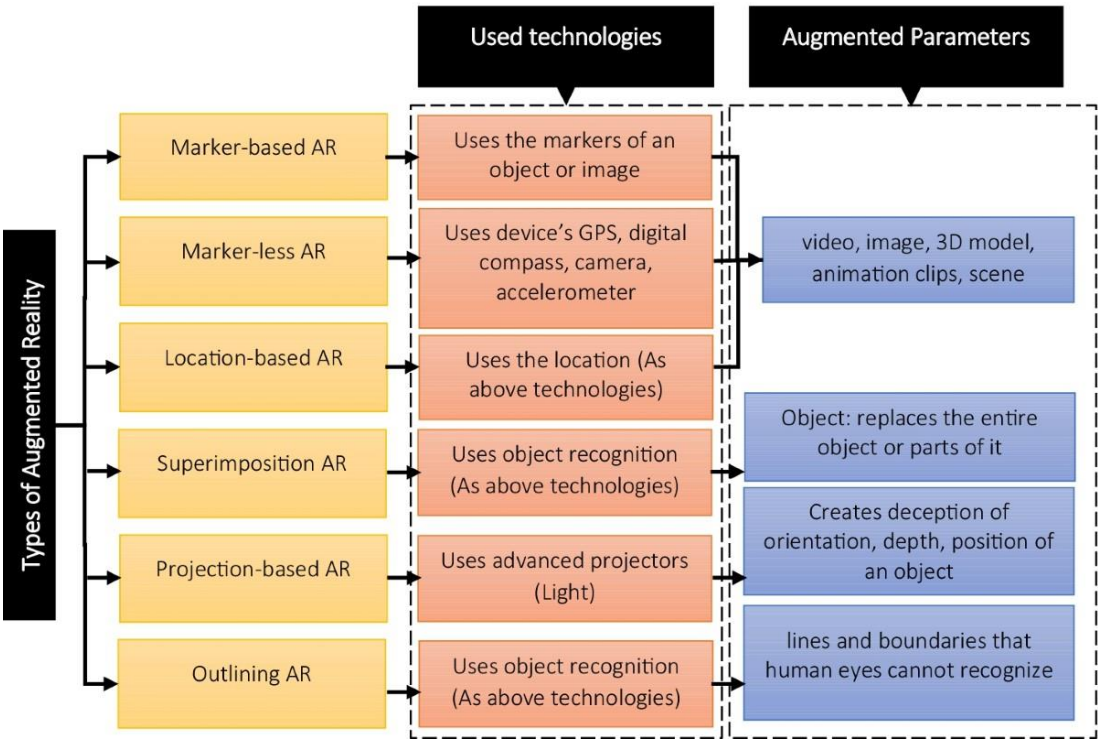

Figure 2. Architectural Components AR & VR

2.4 User Interaction Models

Interaction in AR/VR moves beyond traditional keyboard/mouse or touch interfaces, relying heavily on:
• Spatial Interaction: Users interact with virtual objects positioned in 3D space, using gestures, tracked controllers, or gaze [24]. This requires precise tracking and interpretation of user movements.

• Gaze and Attention: Eye tracking allows systems to know where the user is looking, enabling gazebased selection, foveated rendering, and potentially inferring user interest or cognitive state [25].
• Voice Commands: Natural language interaction adds convenience but introduces risks related to eavesdropping and command injection [26].
• Biometric Input: Continuous collection of motion data (head, hand movements), eye movements, and potentially physiological signals constitutes a rich stream of biometric data, usable for interaction but also vulnerable to misuse.

Table 2. Summary of Key Architectural Components

Reference	Main Focus	Key Contribution/Findings	Limitations/Future Work (as identified)
[14]	Cybersecurity ns implication of eye-tracking	Explores the cybersecurity and privacy risks associated with the use of eye-tracking technology in HMDs, highlighting potential data misuse. [5]	Primarily focused on identifying risks and raising awareness.
[15]	Physiological sensing in XR	Discusses the integration and application of physiological sensors (EEG, EMG, EKG, GSR) in XR systems, likely covering potential uses and associated challenges (including privacy/security).	Emerging field; may focus on specific sensor types or applications.
[16]	Ethical considerations of AR displays	Discusses the ethical implications related to different AR display types (optical see-through vs. video pass-through), potentially including privacy and user perception issues.	Primarily focused on ethics rather than technical security vulnerabilities.
[18]	AR/VR display characteristics and user experience	Focuses on the role of display parameters like resolution and refresh rate in VR immersion and mitigating issues like motion sickness.	Primarily focused on user experience and display technology, not directly security.
[20]	Use of hand tracking via cameras (potentially security implications)	Explores the application of camera based hand tracking in XR, possibly touching upon the associated data privacy and security aspects. (Could be related to De Sutter et al. 2019).	May focus more on the interaction modality than deep security analysis.
[21]	Security aspects of eye tracking for interaction	Focuses on the security and privacy implications of using eye tracking data for gaze-based interaction and techniques like foveated rendering.	Similar focus to Happa et al. (2019), potentially with updated context or different specific threats explored.

[19]	Security vulnerabilities in Metaverse platforms	Investigates security issues within the broader context of the Metaverse, likely covering vulnerabilities in the underlying software and infrastructure, applicable to AR/VR components. [6]	Focuses on the Metaverse concept, which is broader than just AR/VR hardware/software stacks.

Table 3. Summary of Data Flow and Processing

Reference	Main Focus	Key Contribution/Findings	Limitations/Future Work (as identified)
[23]	Sensor fusion algorithms for tracking/SLAM	Likely details methods for combining data from multiple sensors (e.g., IMU, cameras) to achieve robust pose estimation and environmental mapping, critical for AR/VR tracking.	Specific algorithmic limitations or performance trade-offs may be discussed.
[22]	Performance requirements for AR/VR systems	Likely discusses the importance of high frame rates/refresh rates (Hz) for achieving user comfort (preventing motion sickness) and maintaining immersion in AR/VR experiences.	Focus is likely on performance metrics and user experience rather than direct security implications, though system disruption is a security concern.

These interaction models generate vast amounts of sensitive personal data, fundamentally differentiating AR/VR from prior technologies in terms of privacy exposure. Table 4 summarizes the different user interaction models proposed in the literature. The security frameworks must account for the unique nature and sensitivity of this data.

Table 4. Summary of User Interaction Models

Reference	Main Focus	Key Contribution/Findings	Limitations/Future Work (as identified)
[24]	Usability of spatial interaction techniques	Examines the effectiveness and user experience of interacting with 3D environments using methods like gestures, controllers, or gaze.	Focuses on usability rather than security, though poor usability can sometimes lead to errors with security implications.
[25]	(Placeholder/Future Work Citation)	This appears to be a placeholder or speculative citation for future work related to forensics in AR/VR, potentially involving gaze/attention data.	Not a published work, cannot be analyzed.
[26]	Insider threats related to voice command systems	Explores risks associated with voice interactions, possibly focusing on how authorized users (insiders) could misuse the system or be targeted,	May focus on specific system types or threat models.

		including eavesdropping and command injection.	
[27]	Privacy risks of biometric data in sensing systems	Discusses the privacy implications arising from the collection and potential misuse of rich biometric data (movement, physiological signals) gathered by sensor-based systems like AR/VR.	Focuses on identifying and characterizing privacy risks.
[28]	Privacy challenges in the Metaverse/XR	Analyzes the heightened privacy exposure in Metaverse and XR environments due to the collection of vast amounts of sensitive personal and biometric data compared to traditional technologies. [2]	Addresses the broad concept of the Metaverse; specific technical solutions may be outside its scope.

3. Distinct Threat Landscapes: AR vs. VR

While both AR and VR fall under the immersive technology banner, their fundamental difference—AR's augmentation of reality versus VR's replacement of it—creates markedly distinct threat landscapes. Analyzing these differences is crucial for developing targeted security measures. The interaction between the digital overlay and the physical world in AR presents unique challenges, whereas the total sensory immersion of VR opens doors for different forms of manipulation and control.

3.1 The Augmented Reality Threat Surface: Blurring Physical and Digital

The core security challenge in AR stems from its reliance on accurately perceiving and interacting with the physical environment while simultaneously managing digital overlays. The attack surface is characterized by vulnerabilities at this physical-digital interface.
• Sensor Integrity as a Primary Target: AR systems heavily depend on outward-facing sensors (cameras, depth sensors, LiDAR, GPS, and IMUs) to understand the real world for SLAM, object recognition, and correct overlay placement. Malicious manipulation of these sensor inputs, known as sensor spoofing or data injection, is a critical threat. An attacker could feed false camera data, incorrect GPS coordinates, or corrupted IMU readings. The consequences range from misaligned or unstable overlays, causing user confusion or annoyance, to severe safety risks if AR is used for navigation (e.g., in vehicles)
or critical tasks (e.g., surgery), where incorrect information could lead to physical harm. Occlusion attacks, where physical objects block sensor views, can also degrade performance or be exploited.
• Manipulation via AR Overlays: The digital information overlaid onto the real world is a direct channel for potential attacks. Malicious applications or compromised systems could inject false or misleading information, obscure real-world objects or warnings (digital occlusion), or create phantom objects. Consider AR navigation displaying incorrect directions, AR warnings being hidden, or sensitive information being visually leaked through an insecure overlay captured by onlookers or external cameras. Social engineering attacks become potent when delivered through seemingly legitimate AR interfaces. The challenge lies in ensuring the authenticity and integrity of displayed virtual content relative to the physical context [29].
• Contextual Data Exploitation: AR applications often require access to sensitive contextual data, such as the user's precise location, visual data of their surroundings (including private spaces like homes or offices), and recognized objects or people. This data, if compromised, can lead to severe privacy violations. Attackers could reconstruct detailed 3D maps of private environments from leaked SLAM data, infer user activities or relationships based on location and recognized objects, or

engage in targeted surveillance. The permissions model for accessing such data is often coarse-grained, presenting significant risks.

• Physical World Interactions: Since AR users remain aware of and interact with their physical surroundings, attacks can leverage this physical presence. "Man-in-the-Room" (MitR) attacks involve a physically proximate adversary observing the user's interactions with AR content (shoulder surfing virtual screens) or manipulating the environment to interfere with AR system sensors. Physical tampering with AR devices or sensors is also a direct threat vector [30].

3.2 The Virtual Reality Threat Surface: Controlling Simulated Perception

In contrast to AR, VR isolates users within a synthetic world, shifting the primary threat vectors towards manipulation of this controlled sensory experience and exploitation of the data generated within it.

• Sensory Manipulation and Hijacking: The complete control VR systems exert over the user's visual and auditory input makes them susceptible to perception-based attacks. An attacker gaining control of the VR system (VR Hijacking) could alter the virtual environment, mislead the user, induce nausea or disorientation (cyber sickness attacks), inject malicious content (e.g., frightening imagery, unwanted advertisements), or completely deny service by blacking out the display. The psychological impact of such attacks can be significant due to the high level of immersion [31]. Ensuring the integrity of the rendered environment and the communication channels delivering it is paramount.

• Exploitation of Tracking and Interaction Data: VR systems capture highly granular data about user movements – head pose, hand gestures, eye movements, controller positions. This data is essential for interaction but is also a rich source of sensitive information. Unauthorized motion tracking can reveal user activities, identify individuals based on movement patterns (gait, gesture biometrics), infer physical or cognitive states (e.g., fatigue, distraction through eye tracking), or enable avatar puppeteering for impersonation [32]. Logging interactions with virtual objects can reveal user preferences or intentions. The privacy implications of collecting and potentially leaking this data are profound.

• Synthetic Environment Vulnerabilities: The virtual environments themselves can harbor threats. Malicious virtual objects could be designed to trigger unintended actions or extract information when interacted with [33]. Insecure multi-user VR platforms could allow one user to harass, spy on, or attack another user within the virtual space (e.g., virtual groping, avatar exploits). The security of the platform, the communication protocols between users, and the vetting of user-generated content are critical concerns [3].

• Isolation and Reduced Physical Awareness: While immersion is VR's strength, it also makes users vulnerable in the physical world. Users are typically unaware of their immediate physical surroundings, increasing risks of collision or injury. While features like pass through" modes or chaperone systems mitigate this, they can potentially be disabled or spoofed by attackers. Furthermore, the isolation makes users more susceptible to social engineering conducted entirely within the VR environment, as external reality checks are unavailable. Table 5 summarizes the key distinctions in threat landscapes.

Table 5. Comparative Threat Landscape: AR vs. VR

Aspect	Augmented Reality (AR)	Virtual Reality (VR)
Primary Interface	Physical World + Digital Overlay	Fully Synthetic Environment
Key Vulnerability Point	Physical-Digital Interface Integrity	Control over Simulated Perception
Critical Sensors	Outward-facing (Camera, LiDAR, GPS, IMU)	Inward/Tracking (IMU, Tracking Cams, Eye-Track)
Dominant Attack Vectors	Sensor Spoofing, Overlay Manipulation, Contextual Data Leakage, Physical Interference	Environment Hijacking, Sensory Manipulation, Motion/Biometric Data Theft, Synthetic Object Attacks
Primary Impact Domain	Misperception of Real World, Physical Safety Risks, Privacy via Environmental Data	Psychological Manipulation, Virtual Identity Theft, Privacy via Behavioral/Biometric Data, Physical Isolation Risks

Interaction Risk Focus	Insecure Overlay Content, Real-World Distraction/Occlusion	Malicious Virtual Interactions, Avatar Exploits, Social Attacks within VR

4. Emerging Attack Vectors in Immersive Technologies

The unique capabilities and data streams of AR/VR systems enable a range of novel attack vectors beyond traditional cybersecurity threats. These attacks exploit the tight coupling between the digital system, the user's perception, and potentially their physical environment.

4.1 Sensor Spoofing and Data Injection

Manipulating the sensor data that AR/VR systems rely upon poses a fundamental threat to their integrity and safety [34].

• Mechanism: Attackers aim to feed fabricated or altered data into the system's sensor processing pipeline, bypassing or deceiving the actual physical sensors. This can be achieved through various means:

− *Signal Replay/Relay:* Capturing legitimate sensor signals (e.g., GPS, camera feeds) and replaying them later or relaying them from a different location.

− *Direct Injection:* If an attacker gains software access (e.g., via malware), they can directly inject malicious data into the software layers that process sensor input

− *Physical Manipulation:* Using external stimuli to influence sensors, such as projecting infrared patterns to confuse tracking cameras, using acoustic waves to interfere with MEMS IMUs, or employing powerful lights to blind cameras.

• Impact (AR): Successful spoofing in AR can cause virtual objects to appear misaligned, detached from the real world, or unstable, undermining the user experience. More severely, incorrect environmental perception can lead to faulty navigation guidance, failure to recognize hazards, or improper execution of AR-assisted tasks, potentially causing accidents or physical harm. Imagine an AR automotive heads-up display failing to highlight a pedestrian due to spoofed sensor data.

• Impact (VR): While less reliant on external world sensors, VR tracking systems (positional and orientation) are vulnerable. Spoofing IMU data can lead to drift or incorrect orientation, causing disorientation and cyber sickness. Manipulating positional tracking data (e.g., confusing inside-out cameras) could cause the user's virtual viewpoint or avatar position to jump erratically or become mismatched with their physical movements, breaking immersion and potentially leading to collisions with the physical environment if safety boundaries (chaperone systems) are also compromised or based on faulty tracking. Spoofing eye-tracking data could disrupt gaze-based interactions or defeat foveated rendering optimizations. Defending against sensor spoofing requires robust sensor fusion algorithms, anomaly detection, data authentication mechanisms, and potentially hardware-level security features.

4.2 Man-in-the-Room (MitR) Attacks

Distinct from network-based Man-in-the-Middle (MitM) attacks, MitR attacks leverage physical proximity to the AR/VR user [1].

• Mechanism: An adversary situated physically near the user exploits the characteristics of immersive technology usage. This can involve:

− *Shoulder Surfing (Virtual):* Observing the virtual content displayed to an AR user, potentially revealing sensitive information presented on virtual screens or overlays. This is particularly relevant for optical see-through AR where light leakage might occur, or video pass-through if the screen is visible.

− *Environmental Manipulation:* Altering the physical environment to interfere with AR sensors (e.g., placing objects to block cameras, using lights/projectors to disrupt tracking).

− *Eavesdropping:* Capturing audio commands given by the user or conversations held within a VR simulation if microphones are compromised or audio leaks [2].

− *Observation of Physical Actions:* Correlating the user's physical movements (gestures, locomotion) with their actions in the virtual world to infer activities or credentials [1].

• Relevance (AR vs. VR): MitR attacks arguably pose a more significant threat in AR settings, where the user is interacting with virtual content within a shared physical space [35]. In VR, the user's isolation reduces opportunities for virtual shoulder surfing, but physical environment manipulation or observation of movements remains possible, especially if the attacker has access to the same physical space.

• Impact: Information leakage (confidential data, credentials), disruption of service (interfering with tracking), inference of user behavior, and potential for physical annoyance or obstruction.

Mitigation involves user awareness, designing displays to minimize light leakage, robust sensor validation against environmental interference, and potentially context-aware security policies that adjust based on the physical environment's perceived security level.

4.3 Data Leakage through AR Overlays

AR overlays, while enhancing reality, can become conduits for unintentional or malicious data exposure. • Mechanism: Sensitive information can be leaked through AR displays in several ways:

− *Inadvertent Display:* Applications might display private information (e.g., messages, notifications, medical data) as an overlay that is visible to nearby observers (MitR) or captured by ambient cameras

− *Malicious Overlays:* A malicious application could generate overlays designed to trick the user into revealing information or overlays that secretly capture portions of the user's view (potentially containing sensitive real-world details) and exfiltrate them

− *Overlay Content Inference:* Even if the overlay content itself isn't directly sensitive, its placement, timing, or type could allow an observer or a compromised system component to infer the user's activity, context, or focus of attention

− *Screen Recording/Streaming:* Unauthorized capturing of the AR view (including both real world and overlays) through screen recording malware or insecure streaming protocols

• Impact: Compromise of personal privacy, leakage of confidential business or medical information, enabling social engineering attacks based on leaked context. The persistence of AR overlays means that sensitive information might be displayed continuously in the user's field of view.

• Challenge: Balancing the utility of readily available AR information with the need to protect sensitive data visible within the overlay, particularly in shared or public spaces.

Mitigation strategies include context-aware display policies (e.g., dimming or hiding sensitive overlays based on location or presence of others), secure rendering pipelines that prevent unauthorized capture, fine-grained permissions for applications requesting overlay privileges, and potentially privacy-preserving rendering techniques that obfuscate sensitive regions [36].

4.4 VR Hijacking and Perception Manipulation

Gaining control over a user's immersive VR experience represents a powerful attack vector with potentially severe psychological consequences.

• Mechanism: An attacker compromises the VR system (host PC, console, standalone HMD OS, or network connection) to manipulate the sensory output delivered to the user. This could involve:

− *Visual Manipulation:* Altering rendered frames to change the appearance of the virtual environment, insert unexpected or malicious objects/characters, display false information, or replace the entire scene

− *Auditory Manipulation:* Injecting false sounds, altering spatial audio cues to mislead the user, or eavesdropping/recording microphone input

− *Haptic Manipulation:* Triggering unexpected haptic feedback or disabling expected feedback, potentially causing discomfort or confusion [37].

− *Denial of Service:* Preventing rendering entirely (black screen), inducing extreme cybersickness through mismatched motion cues, or locking the user out of controls.

• Impact: Disorientation, fear, psychological distress, potential for inducing specific behaviors through manipulated cues (e.g., guiding a user towards a physical obstacle), delivery of propaganda or unwanted advertising, erosion of trust in the VR system. The high level of immersion can make users particularly susceptible to manipulation.

- Attack Surface: Vulnerabilities in the OS, drivers, rendering pipeline, application software, and network protocols used for remote rendering or multi-user experiences. Standalone HMDs running mobile OS variants inherit mobile security challenges

Defenses require secure boot processes, OS hardening, integrity protection for the rendering pipeline (e.g., using Trusted Execution Environ), secure network protocols, application vetting, and anomaly detection systems monitoring for unexpected sensory output patterns.

4.5 Unauthorized Motion Tracking and Behavioral Inference

AR/VR systems intrinsically capture detailed motion data for interaction, creating a significant privacy risk if accessed or analyzed illegitimately.

- Mechanism: Tracking data (head position/orientation, hand/controller movements, eye movements, body posture if tracked) is collected continuously at high frequency. An attacker gaining access to this data stream (e.g., through malware, network eavesdropping, compromised application/SDK) can perform various analyses:

– *Activity Recognition:* Inferring user actions (e.g., typing on a virtual keyboard, specific gestures, walking patterns).

– *Biometric Identification/Authentication Bypass:* Using unique patterns in head movement, gait, or gestures as biometric identifiers to re-identify users across sessions or platforms, or potentially to impersonate them [38].

– *Health/Cognitive State Inference:* Analyzing motion patterns (e.g., tremors, reaction times, eye movement characteristics like saccades and pupil dilation) to infer potential health conditions (neurological disorders), fatigue, intoxication, or cognitive load.

– *Skill/Intention Inference:* Analyzing fine motor control during tasks to infer skill level or predicting user intentions based on preliminary movements or gaze patterns.

- Impact: Profound privacy loss, enabling unwanted profiling, discrimination based on inferred health or behavior, potential for blackmail, and undermining user anonymity. The granularity and continuous nature of this data make it exceptionally revealing.
- Challenge: Balancing the need for high-fidelity tracking for seamless interaction with the imperative to protect users from the misuse of this highly personal data. Current privacy policies and consent mechanisms often fail to adequately convey the extent and sensitivity of motion data collection.

Mitigation approaches include data minimization (collecting only necessary data), on-device processing, and differential privacy techniques for aggregated data analysis. Cryptographic methods, access control mechanisms limiting data access for applications, and user-controlled obfuscation or noise injection features.

5. Cross-cutting security and privacy challenges

Beyond specific attack vectors tied to AR or VR architectures, several security and privacy challenges cut across both technologies, amplified by the immersive and data-rich nature of these systems. These issues often relate to the intimate user data collected and the social dynamics within virtual spaces.

5.1 Biometric Data Misuse

AR/VR systems are inherently biometric data collection devices, capturing information far beyond traditional systems [39].
• Types of Biometric Data:
– Behavioral Biometrics: Continuous head movement patterns, hand/arm gesture dynamics, gait patterns from locomotion tracking, interaction patterns with virtual objects.
– Physiological Biometrics: Eye-tracking data (pupil dilation, saccades, fixations, vengeance). Voice patterns from microphone input. Potentially EEG, EKG, and GSR if specialized sensors are integrated • Risks:
– Unauthorized Identification/Tracking: Using unique biometric signatures to identify users across different applications, platforms, or sessions, even when attempting to remain anonymous. The stability and uniqueness of some XR biometrics are active research areas.
– Inference of Sensitive Attributes: As discussed previously, biometric data can be used to infer health status, emotional state, cognitive load, fatigue, substance influence, or even sexual orientation (e.g., inferred from pupil dilation responses [40]) [41]. This raises significant ethical concerns [42].

– Authentication System Vulnerabilities: If biometrics are used for authentication (e.g., gaze patterns to unlock a device), they become targets for spoofing or replay attacks. Unlike passwords, compromised biometrics cannot easily be changed.

– Function Creep: Biometric data collected for one purpose (e.g., foveated rendering) could later be repurposed for other uses (e.g., emotion tracking for advertising) without explicit user consent.

• Challenges: The sheer volume and continuous nature of biometric data collection in XR. Obtaining meaningful user consent when the implications are complex and evolving. Secure storage and processing of highly sensitive biometric data. Lack of standards and regulations specifically addressing XR biometrics.

Addressing these risks requires strong privacy-preserving techniques, robust security architectures, transparent policies, user control over data collection, and potentially new regulatory frameworks.

5.2 Identity Theft and Misrepresentation in Virtual Spaces

As social interactions become common in shared VR and AR environments (the "Metaverse" concept), issues of identity management, theft, and misrepresentation become critical.

• Mechanisms:

– Avatar Theft/Impersonation: Gaining unauthorized access to a user's account and controlling their avatar to interact with others, potentially damaging their reputation, extracting information from contacts, or conducting scams. The visual and behavioral fidelity of avatars can make impersonation convincing.

– Avatar Hijacking/Puppeteering: Temporarily taking control of another user's avatar movements or voice output through system exploits, potentially forcing them to perform unwanted actions.

– Deep fakes in VR/AR: Using AI to generate realistic but fake representations of individuals (avatars, voice) within immersive environments. This could be used for sophisticated impersonation, spreading misinformation, or harassment. Generating convincing real-time deep fakes presents technical challenges but is an evolving threat.

– Misleading Avatars: Users intentionally creating avatars that misrepresent their identity or intentions to deceive or manipulate others in social VR/AR settings.

– Linking Virtual and Real Identities: Exploiting data leaks or system vulnerabilities to link a user's seemingly anonymous virtual persona to their real-world identity, exposing them to doxing or harassment.

• Impact: Financial loss (scams), reputational damage, social engineering, psychological distress from harassment or impersonation, erosion of trust in virtual platforms, spread of misinformation. The persistence and social nature of virtual worlds can amplify the impact of identity-related attacks.

• Challenges: Secure authentication and session management for VR/AR accounts. Verifying identity in virtual spaces without sacrificing privacy. Detecting and mitigating deep fakes and sophisticated impersonation attempts. Establishing and enforcing codes of conduct in large-scale virtual environments. Balancing freedom of expression with the need to prevent malicious misrepresentation.

Solutions involve robust multi-factor authentication, secure avatar/profile management systems, potentially verifiable credentials for certain interactions, AI-based detection of anomalous behavior or deep fakes, platform moderation, and user education [43].

5.3 Spatial-Temporal Data Inference

Both AR and VR systems generate data that reveals information about the user's physical environment and movement patterns over time, posing significant privacy risks.

• Data Sources:

– SLAM Data (AR): The maps of the physical environment created by AR systems for tracking and overlay placement contain detailed geometric information about rooms, objects, and potentially sensitive locations (homes, workplaces).

– Location Data (AR/VR): GPS/GNSS data (primarily AR), Wi-Fi/Bluetooth-based localization, and potentially inferred location based on room mapping or chaperone boundary setup (VR).

– Movement Trajectories (AR/VR): Paths taken by the user within a physical space (AR locomotion, VR room-scale movement) or within large virtual environments.

– Interaction Hotspots: Areas in the physical (AR) or virtual (VR) world where the user frequently interacts or gazes.

• Inference Risks:

– Environmental Reconstruction: Leaked SLAM data can allow attackers to create detailed 3D models of private spaces, revealing layouts, valuable objects, or security systems.

– Routine and Habit Inference: Analyzing movement patterns over time (spatial-temporal data) can reveal daily routines, places visited, work schedules, social connections (if co-location data is available).

– Presence Detection: Knowing when a user is active in a particular location (e.g., home vs. office).

– Object/Activity Inference: Correlating location/map data with recognized objects or interaction patterns to infer user activities, interests, or socioeconomic status.
• Challenges: The necessity of spatial mapping for core AR functionality and room-scale VR. The difficulty in anonymizing detailed geometric map data. Users' potential lack of awareness regarding the extent of environmental data being collected and stored. The tension between providing personalized, context-aware experiences and protecting spatial privacy. Mitigation requires careful data management practices, including on-device storage and processing where possible, data minimization (e.g., storing only features needed for tracking, not full maps), map data encryption, techniques for privacy-preserving spatial data sharing (e.g., secure multi-party computation for collaborative AR), user controls over map storage and sharing, and transparent data usage policies.

6. Mitigation strategies and defense mechanisms

Addressing the diverse security and privacy threats in AR and VR necessitates a multi-layered approach, combining hardware security, secure software engineering practices, cryptographic techniques, advanced algorithms, and user-centric controls. No single solution is sufficient; resilience requires integrating defenses across the entire system architecture.

6.1 Secure Sensor Integration and Fusion

Given the criticality of sensor data, ensuring its integrity and trustworthiness is paramount, especially for AR.
• Hardware Security Modules (HSMs) and Trusted Execution Environments (TEEs): Integrating sensors with HSMs or processing sensitive sensor data within TEEs (like ARM Trust Zone or Intel SGX) can provide hardware-level isolation and protection. This can help ensure data integrity, confidentiality, and authentic origin, making software-based injection attacks more difficult. However, TEEs themselves can have vulnerabilities and performance overhead.
• Sensor Data Authentication: Implementing protocols to authenticate sensor readings, perhaps using physical unclonable functions (PUFs) embedded near sensors or cryptographic signatures applied close to the source, can help detect tampering or replay attacks. Challenges include key management and performance impact.
• Robust Sensor Fusion Algorithms: Designing sensor fusion algorithms (e.g., combining IMU, camera, and GPS data) that are resilient to outliers or failures in individual sensor streams. Techniques like Kalman filters or particle filters can incorporate uncertainty modeling. Cross-validation between different sensor modalities can help detect inconsistencies indicative of spoofing.
• Anomaly Detection: Employing machine learning models or statistical methods to monitor sensor data streams and detect deviations from expected patterns or physical plausibility. For instance, detecting physically impossible movements or inconsistencies between visual odometer and IMU readings. These systems need careful tuning to avoid false positives.
• Physical Tamper Detection: Incorporating mechanisms to detect physical tampering with sensor hardware or enclosures.

6.2 Encryption and Secure Communication

Protecting data in transit and at rest is a fundamental security requirement, especially given the volume and sensitivity of AR/VR data streams [44].
• End-to-End Encryption (E2EE): Applying E2EE to all sensitive data streams, including video/audio feeds, sensor data, tracking information, and communication between users in multi-user environments [45]. This ensures that intermediate nodes (e.g., cloud servers used for rendering or computation) cannot access the plaintext data. Standard protocols like TLS/DTLS and secure messaging protocols (e.g., Signal protocol can be adapted. Challenges include key management and potential performance overhead impacting real-time requirements.
• Data-at-Rest Encryption: Encrypting sensitive data stored on the device (HMD, PC, mobile), such as biometric profiles, SLAM maps, user preferences, and cached application data. Utilizing hardware backed key stores available in modern processors and secure elements enhances protection.
• Secure Protocols for Device/Cloud Interaction: When offloading computation or rendering to edge/cloud servers, secure and authenticated communication channels are essential. This includes mutual authentication between device and server, integrity protection for transmitted data, and protection against replay attacks.
• Homomorphic Encryption and Secure Multi-Party Computation (SMC): For privacy-sensitive computations involving multiple parties or cloud processing, advanced cryptographic techniques like homomorphic encryption (allowing computation on encrypted data) or SMC (allowing multiple parties to compute a function on their private inputs without revealing them) offer potential, though often with significant performance costs currently limiting real-time application [46]. Research is ongoing to make these practical for XR workloads [47].

6.3 Context-Aware Access Control

Standard permission models (e.g., grant access to camera at install time) are often too coarse-grained for the sensitive data and capabilities of AR/VR systems. Context-aware access control offers a more nuanced approach.
• Fine-Grained Permissions: Breaking down permissions into more specific capabilities (e.g., access raw camera feed vs. access object recognition results, access coarse location vs. precise location, access head orientation vs. full 6DoF pose).
• Dynamic Policy Enforcement: Granting or restricting access based on the current context, such as:
– Location: Restricting access to sensitive data (e.g., SLAM maps of home) when the user is in a public space.
– Activity: Allowing sensor access only when the application is actively performing a relevant task.
– Presence of Others: Automatically obfuscating sensitive information on AR overlays if other people are detected nearby.
– Application State: Limiting background data access for applications not in the foreground.
• User Control and Transparency: Providing users with clear information about what data is being accessed, by which applications, and for what purpose, along with intuitive controls to manage these permissions. Privacy dashboards and just-in-time prompts can improve transparency.
• Risk-Based Access Control: Adapting access based on the trustworthiness of the application, the sensitivity of the data requested, and the potential risk associated with granting access in the current context [48].
Implementing effective context-aware access control requires reliable context sensing and robust policy definition and enforcement mechanisms within the AR/VR operating system or middleware [49].

6.4 Privacy-Preserving Rendering and Computation

Techniques aimed at processing or displaying sensitive information in ways that minimize privacy leakage are crucial, particularly for visual data.
• Privacy-Preserving Rendering (AR): Techniques to prevent unintended information leakage from AR overlays:
– Visual Obfuscation/Redaction: Automatically blurring, pixelating, or redacting sensitive regions of the real-world view captured by cameras (e.g., faces of bystanders, license plates, content on screens) before processing or display in video pass-through AR [50]. Similar techniques can be applied to sensitive overlay content [21].
– Content-Aware Display Adaptation: Modifying how information is displayed based on sensitivity and context (e.g., showing only essential information in public, using abstract representations)
[4].
– Secure Rendering Paths: Utilizing TEEs or dedicated hardware paths to render sensitive overlays directly to the display, preventing intermediate software (including malware or screen recorders) from capturing them.
• Differential Privacy: Adding carefully calibrated noise to data (e.g., aggregated motion statistics, location histograms) before sharing or analysis, providing mathematical guarantees that individual user contributions cannot be easily identified. Applicable to analytics derived from AR/VR usage but less suitable for real-time interaction data requiring high fidelity. Research explores applying DP concepts to trajectory or SLAM data.
• Federated Learning: Training machine learning models (e.g., for activity recognition, personalization) directly on user devices without sending raw data to a central server. Only model updates are aggregated centrally, improving privacy. Applicable to various XR tasks involving ML
• Data Minimization and Anonymization: Collecting and retaining only the minimum data necessary for functionality. Applying anonymization techniques where possible, although robust anonymization of high-dimensional data like motion or SLAM maps remains challenging. Techniques like k-anonymity or l-diversity may offer limited protection.

6.5 User Awareness, Training, and Secure Development

Technical solutions must be complemented by human factors and secure development practices.
• User Education: Informing users about the potential security and privacy risks associated with AR/VR, safe usage practices (e.g., being aware of surroundings in VR, managing permissions), and how to recognize potential attacks (e.g., social engineering in VR).
• Secure Software Development Lifecycle (SSDLC): Integrating security considerations throughout the development process for AR/VR hardware, OS, SDKs, and applications. This includes threat modeling specific to AR/VR. Secure coding practices, vulnerability testing (static/dynamic analysis, fuzzing), and secure update mechanisms.
• Application Vetting: Rigorous review processes for AR/VR applications distributed through app stores to identify malware, excessive permission requests, or privacy violations. Implementing these mitigation strategies effectively requires a concerted effort from hardware manufacturers, platform developers, application creators, researchers, and policymakers. The dynamic nature of threats necessitates continuous adaptation and improvement of defenses.

7. Research gaps and future directions

Despite progress in understanding and mitigating AR/VR security and privacy risks, significant research gaps remain. The rapid evolution of immersive technologies continually introduces new challenges, demanding ongoing investigation and innovation to ensure trustworthy systems. This section outlines key areas requiring further research and proposes a roadmap for future work.

7.1 Identified Research Gaps

• Lack of Comprehensive Security Standards and Benchmarks: Currently, there are few widely accepted security standards specifically tailored to AR/VR hardware, software platforms, or applications [51]. This hinders interoperability, consistent security assessment, and the development of baseline security requirements. Establishing common criteria and benchmarks for evaluating the security posture of AR/VR systems is critical.
• Scalability and Performance of Security Mechanisms: Many proposed security and privacy techniques (e.g., advanced cryptography like HE/SMC, complex TEE usage, sophisticated anomaly detection) incur significant computational or latency overhead [52]. Given the stringent real-time performance requirements of immersive experiences (high frame rates, low latency), developing lightweight yet effective security mechanisms that scale to complex applications and multi-user environments remains a major challenge [2].
• Usability vs. Security and Privacy Trade-offs: Implementing robust security measures often introduces friction into the user experience (e.g., frequent permission prompts, complex configuration settings). Finding the right balance between strong protection and seamless, intuitive interaction is crucial for user acceptance [53]. Research is needed into usable security and privacy controls designed specifically for the interaction paradigms of AR/VR (e.g., gaze-based or gesture-based consent mechanisms) [54].
• Longitudinal Privacy Protection: AR/VR systems have the potential to collect vast amounts of sensitive data over extended periods, creating detailed longitudinal records of user behavior, environments, and biometrics [55]. Current privacy frameworks often focus on transactional data protection. Developing methods to protect privacy against inferences drawn from long-term data aggregation, manage data deletion securely (the "right to be forgotten" [56]), and mitigate risks from evolving analysis techniques is an open problem.
• Cross-Platform and Interoperability Security: As users interact across different AR/VR devices, platforms, and applications (potentially forming parts of a "Metaverse"), ensuring consistent security and privacy policies, secure data transfer between environments, and managing identities across platforms becomes increasingly complex [5]. Addressing vulnerabilities arising from interactions between heterogeneous systems is essential.
• Security of AI/ML Models in AR/VR: Immersive systems heavily rely on machine learning for crucial functions like SLAM, object recognition, hand/eye tracking, and behavioral analysis. These ML models are themselves vulnerable to adversarial attacks (e.g., evasion attacks, data poisoning, model inversion). Research into robust and verifiable AI specifically for the AR/VR context, including detecting and mitigating adversarial inputs targeting perception or tracking algorithms, is needed. Understanding the privacy implications of deployed ML models (e.g., inferring unintended information) is also critical.
• AR/VR Forensics and Incident Response: Investigating security incidents within immersive environments presents unique challenges. Collecting relevant evidence from ephemeral AR overlays or complex VR interactions, analyzing large volumes of sensor and interaction data, and attributing actions within virtual spaces require new forensic tools and methodologies. Developing frameworks for incident response tailored to AR/VR attacks is necessary.
• Ethical Frameworks and Governance: Beyond technical solutions, clear ethical guidelines and governance structures are needed to address the profound societal implications of AR/VR. This includes issues of data ownership, algorithmic bias in perception/analysis, potential for manipulation or addiction, equitable access, and the impact on social interaction and reality perception. Establishing norms and potentially regulations for responsible development and deployment is crucial.
• Security of Novel Interaction Modalities: As new interaction methods emerge (e.g., Brain-Computer Interfaces (BCIs), advanced haptics), understanding and mitigating their unique security and privacy implications (e.g., leakage of neural signals) will be essential.

7.2 Roadmap for Future Research and Development

Addressing these gaps requires a multi-pronged approach: 1. Develop Foundational Security Primitives: Focus on creating efficient cryptographic protocols, lightweight TEE/HSM solutions, and robust sensor authentication mechanisms specifically optimized for the resource constraints and real-time demands of AR/VR. 2. Advance Privacy-Enhancing

Technologies (PETs): Investigate and refine PETs like differential privacy, federated learning, homomorphic encryption, and secure multi-party computation for practical application to AR/VR data types (motion, spatial maps, and biometrics). Develop novel techniques for privacy-preserving rendering and data sharing. 3. Design Resilient Perception and Tracking Algorithms: Create SLAM, tracking, and recognition algorithms that are inherently more robust to sensor noise, spoofing attempts, and adversarial perturbations. Explore verifiable computing concepts for perception pipelines. 4. Build Secure and Usable Platforms: Develop AR/VR operating systems and middleware with built-in, fine-grained, context-aware security and privacy controls. Design user interfaces for managing these controls that are intuitive within immersive interaction paradigms. 5. Establish Standards and Best Practices: Foster collaboration between industry, academia, and standards bodies (like IEEE, ISO, Khronos Group [57]) to develop security and privacy standards, testing methodologies, and best practice guidelines for AR/VR development. 6. Promote Secure Development Culture: Encourage the adoption of SSDLC practices, provide developers with tools and training for building secure AR/VR applications, and foster transparency through responsible vulnerability disclosure programs. 7. Investigate Long-Term Implications: Conduct longitudinal studies on the privacy and security impacts of prolonged AR/VR use. Develop technical and policy solutions for managing long-term data retention and inference risks. 8. Foster Interdisciplinary Research: Encourage collaboration between computer scientists, engineers, psychologists, ethicists, legal scholars, and policymakers to address the complex socio-technical challenges of AR/VR security and privacy. Develop comprehensive ethical frameworks and governance models. 9. Develop Forensics Capabilities: Create specialized tools and techniques for digital forensics and incident response in AR/VR environments.

By pursuing this roadmap, the research community and industry can work towards building a future where the transformative potential of immersive technologies can be realized safely and responsibly, fostering user trust and enabling widespread adoption.

8. Conclusion and Future work

Augmented and Virtual Reality technologies stand at the cusp of widespread adoption, promising unprecedented levels of interaction and immersion that could reshape numerous aspects of modern life. This review has underscored that alongside this immense potential lies a complex and evolving landscape of security and privacy challenges, distinct from those encountered in traditional computing. The intimate coupling of these systems with user perception, motion, biometrics, and often the physical environment creates novel attack surfaces and amplifies the potential consequences of breaches.

We conducted a comparative analysis contrasting the threat landscapes of AR and VR, highlighting how AR's blending of physical and digital realities makes sensor integrity and overlay security paramount, while VR's complete sensory immersion makes it susceptible to perceptual manipulation and the exploitation of rich interaction data. Emerging attack vectors, including sophisticated sensor spoofing, insidious Man-in the-Room tactics, AR overlay data leakage, perception-altering VR hijacking, and unauthorized tracking for behavioral inference, were examined in detail. Furthermore, cross-cutting issues such as the pervasive collection and potential misuse of biometric data, the risks of identity theft within increasingly social virtual spaces, and privacy threats arising from spatial-temporal data analysis were explored.

In response to these threats, a range of mitigation strategies were evaluated. Secure sensor integration using hardware and algorithmic robustness, comprehensive encryption of data in transit and at rest, context aware access control mechanisms, and various privacy-preserving computation and rendering techniques form key pillars of defense. However, the effective implementation of these strategies faces challenges related to performance overhead, usability, and the lack of standardization.

Crucially, significant research gaps persist. Addressing the need for tailored security standards, scalable and efficient defenses, usable privacy controls, longitudinal data protection, cross-platform security assurances, and AI/ML robustness in XR, forensic capabilities, and comprehensive ethical frameworks is imperative. The proposed roadmap emphasizes foundational security advancements, progress in PETs, resilient system design, standardization efforts, secure development practices, interdisciplinary collaboration, and a focus on long-term implications.

The journey towards secure and trustworthy immersive technologies requires a proactive, collaborative, and continuous effort from all stakeholders. Failing to adequately address the security and privacy challenges outlined herein could not only expose users to significant harm—ranging from data theft and psychological manipulation to potential physical danger—but also impede the societal acceptance and ultimate success of AR and VR. By prioritizing security and privacy from the outset, integrating robust defenses across the technology stack, and fostering ongoing research and dialogue, we can strive to build an immersive future that is both innovative and safe.

Corresponding author
Qais Al-Na'amneh
q_naamneh@asu.edu.jo

Contributions
M.A; Y.B; Q.A; M.A.L; R.H; S.M.S; Conceptualization, M.A; Y.B; Q.A; M.A.L; R.H; S.M.S; Investigation, M.A; Y.B; Q.A; M.A.L; R.H; S.M.S; Writing (Original Draft), M.A; Y.B; Q.A; M.A.L; R.H; S.M.S; Writing (Review and Editing) Supervision, M.A; Y.B; Q.A; M.A.L; R.H; S.M.S; Project Administration.

Ethics declarations
This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication
Not applicable.

Competing interests
The author declares no competing interests.

References

[1] Jaradat, A. S., Nasayreh, A., Al-Na'amneh, Q., Gharaibeh, H., & Al Mamlook, R. E. (2023). Genetic optimization techniques for enhancing web attacks classification in machine learning. In *2023 IEEE International Conference on Dependable, Autonomic and Secure Computing, Pervasive Intelligence and Computing, Cloud and Big Data Computing, Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)* (pp. 0130–0136). IEEE.

[2] Aljaidi, M., Alsarhan, A., Al-Fraihat, D., Al-Arjan, A., Igried, B., El-Salhi, S. M., Khalid, M., & Al-Na'amneh, Q. (2023). Cybersecurity threats in the era of AI: Detection of phishing domains through classification rules. In *2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)* (pp. 1–6). IEEE.

[3] Al-Na'amneh, Q., Almaiah, M. A., Smadi, S., Hazaymih, R., & Alabadi, L. (2025). Attacks detection and mitigation of IoT using machine learning model. In *Utilizing AI in Network and Mobile Security for Threat Detection and Prevention* (pp. 115–132). IGI Global Scientific Publishing.

[4] Al-Na'amneh, Q., Aljaidi, M., Nasayreh, A., Gharaibeh, H., Al Mamlook, R. E., Jaradat, A. S., Alsarhan, A., & Samara, G. (2024). Enhancing IoT device security: CNN-SVM hybrid approach for real-time detection of DoS and DDoS attacks. *Journal of Intelligent Systems*.

[5] Alshuaibi, A., Almaayah, M., & Ali, A. (2025). Machine learning for cybersecurity issues: A systematic review. *Journal of Cyber Security and Risk Auditing, 2025*(1), 36–46.

[6] Alghamdi, A., Alkinoon, A., Alghuried, A., & Mohaisen, D. (2024). XR-Droid: A benchmark dataset for AR/VR and security applications. *IEEE Transactions on Dependable and Secure Computing*.

[7] Roesner, F., & Kohno, T. (2021). Security and privacy for augmented reality: Our 10-year retrospective. In *VR4Sec: 1st International Workshop on Security for XR and XR for Security*.

[8] De Guzman, J. A., Thilakarathna, K., & Seneviratne, A. (2019). Security and privacy approaches in mixed reality: A literature survey. *ACM Computing Surveys (CSUR), 52*(6), 1–37.

[9] Lebeck, K., Ruth, K., Kohno, T., & Roesner, F. (2018). Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)* (pp. 392–408). IEEE.

[10] Li, X., Yi, W., Chi, H.-L., Wang, X., & Chan, A. P. (2018). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in Construction, 86*, 150–162.

[11] Kachur, A., Lysenko, S., Bodnaruk, O., & Gaj, P. (2024). Methods of improving security and resilience of VR systems' architecture. In *IntelITSIS* (pp. 285–304).

[12] Mushtaq, M., Jha, R., Sabraj, M., & Jain, S. (2024). Attack modeling and security analysis using machine learning algorithms enabled with augmented reality and virtual reality. *International Journal of Computer Networks and Information Security (IJCNIS)*, 16(4), 118–129.

[13] Gulhane, A., Vyas, A., Mitra, R., Oruche, R., Hoefer, G., Valluripally, S., Calyam, P., & Hoque, K. A. (2019). Security, privacy and safety risk assessment for virtual reality learning environment applications. In *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)* (pp. 1–9). IEEE.

[14] Happa, J., Glencross, M., & Steed, A. (2019). Cyber security threats and challenges in collaborative mixed-reality. *Frontiers in ICT*, 6, 5.

[15] Zhang, Y., Slocum, C., Chen, J., & Abu-Ghazaleh, N. (2023). It's all in your head(set): Side-channel attacks on AR/VR systems. In *32nd USENIX Security Symposium (USENIX Security 23)* (pp. 3979–3996).

[16] Adams, D., Bah, A., Barwulor, C., Musaby, N., Pitkin, K., & Redmiles, E. M. (2018). Ethics emerging: The story of privacy and security perceptions in virtual reality. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)* (pp. 427–442).

[17] Giaretta, A. (2024). Security and privacy in virtual reality: A literature survey. *Virtual Reality*, 29(1), 10.

[18] Milon, A. G., & tom Dieck, M. (2024). AR and VR in the spotlight: A systematic literature review of security, privacy, and ethical concerns. In *Smart Ethics in the Digital World: Proceedings of the ETHICOMP 2024* (pp. 54–57). Universidad de La Rioja.

[19] Chen, Z., Wu, J., Gan, W., & Qi, Z. (2022). Metaverse security and privacy: An overview. In *2022 IEEE International Conference on Big Data (Big Data)* (pp. 2950–2959). IEEE.

[20] de Armas, C., Tori, R., & Netto, A. V. (2020). Use of virtual reality simulators for training programs in the areas of security and defense: A systematic review. *Multimedia Tools and Applications*, 79(5), 3495–3515.

[21] Hoole, R., & Jahankhani, H. (2021). Security framework for delivery of training, using VR technology. In *Information Security Technologies for Controlling Pandemics* (pp. 357–386).

[22] Bang, J., Lee, Y., Lee, Y.-T., & Park, W. (2019). AR/VR based smart policing for fast response to crimes in safe city. *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 470–475. IEEE.

[23] Bhattacharya, P., Saraswat, D., Dave, A., Acharya, M., Tanwar, S., Sharma, G., & Davidson, I. E. (2021). Coalition of 6G and blockchain in AR/VR space: Challenges and future directions. *IEEE Access, 9*, 168455–168484.

[24] Datcu, D., Cidota, M., Lukosch, H., & Lukosch, S. (2014). On the usability of augmented reality for information exchange in teams from the security domain. *2014 IEEE Joint Intelligence and Security Informatics Conference*, 160–167. IEEE.

[25] AR, A. R. (2025). Forensic challenges in augmented reality (AR) and virtual reality (VR).

[26] Choi, T. (2021). Insider threats and lessons learned from the US for South Korea and the use of new technologies like VR/AR to enhance the cybersecurity of embassies and national security (Master's thesis, San Diego State University).

[27] Mustafa, T., Matovu, R., Serwadda, A., & Muirhead, N. (2018). Unsure how to authenticate on your VR headset? Come on, use your head! In *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics* (pp. 23–30).

[28] Zhao, R., Zhang, Y., Zhu, Y., Lan, R., & Hua, Z. (2023). Metaverse: Security and privacy concerns. *Journal of Metaverse, 3*(2), 93–99.

[29] Zhang, S. (2025). Internet of multimodality: Problems in security, healthcare and AR/VR.

[30] Chaudhari, A., Mali, Y. K., Kulkarni, A., Jain, D., Sharma, L., Mahajan, K., Kazi, F., Kar, P., & Bhogle, A. (2024). Cyber security challenges in social meta-verse and mitigation techniques. *2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon)*, 1–7. IEEE.

[31] Ali, S., Abdullah, T. P. T., Armand, A., Athar, A., Hussain, A., Ali, M., Yaseen, M., Joo, M.-I., & Kim, H.-C. (2023). Metaverse in healthcare integrated with explainable AI and blockchain: Enabling immersiveness, ensuring trust, and providing patient data security. *Sensors, 23*(2), 565.

[32] Aljawarneh, M. (2025). Ethical issues in cyber-security for autonomous vehicles (AV) and automated driving. In *Utilizing AI in Network and Mobile Security for Threat Detection and Prevention* (p. 173). IGI Global.

[33] Laila, D. A., Al-Na'amneh, Q., Aljaidi, M., Nasayreh, A. N., Gharaibeh, H., Al Mamlook, R., & Alshammari, M. (2024). Enhancing 2D logistic chaotic map for gray image encryption. In *Risk Assessment and Countermeasures for Cybersecurity* (pp. 170–188). IGI Global.

[34] Truong, V. T., & Le, L. B. (2024). Security for the metaverse: Blockchain and machine learning techniques for intrusion detection. *IEEE Network, 38*(5), 204–212.

[35] Svedberg, J., & Olsson, T. (2018). Augmenting security systems—the role of augmented reality in the surveillance industry.

[36] Raybourn, E. M., & Trechter, R. (2018). Applying model-based situational awareness and augmented reality to next-generation physical security systems. In *Cyber-Physical Systems Security* (pp. 331–344). Springer.

[37] Kumari, S., & Polke, N. (2018). Implementation issues of augmented reality and virtual reality: A survey. In *International Conference on Intelligent Data Communication Technologies and Internet of Things* (pp. 853–861). Springer.

[38] Bhalla, A., Sluganovic, I., Krawiecka, K., & Martinovic, I. (2021). MoveAR: Continuous biometric authentication for augmented reality headsets. In *Proceedings of the 7th ACM on Cyber-Physical System Security Workshop* (pp. 41–52).

[39] Elkoubaiti, H., & Mrabet, R. (2018). How are augmented and virtual reality used in smart classrooms? In *Proceedings of the 2nd International Conference on Smart Digital Environment* (pp. 189–196).

[40] Valles, L., Moriana, A., & Garcia, R. (2018). The TARGET project: Using VR and AR to improve police training.

[41] Aukstakalnis, S. (2016). *Practical augmented reality: A guide to the technologies, applications, and human factors for AR and VR*. Addison-Wesley Professional.

[42] Park, S., Kim, J. W., Kim, K. M., & Kim, H. (2018). AR-based field training system algorithm for small units. *Convergence Security Journal, 18*(4), 81–87.

[43] Carneiro, J., Rossetti, R. J., Silva, D. C., & Oliveira, E. C. (2018). BIM, GIS, IoT, and AR/VR integration for smart maintenance and management of road networks: A review. In *2018 IEEE International Smart Cities Conference (ISC2)* (pp. 1–7). IEEE.

[44] Kittinger, L. I. (2018). VR proposal for breakout session. Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), Tech. Rep.

[45] Meyer-Lee, G., Shang, J., & Wu, J. (2018). Location-leaking through network traffic in mobile augmented reality applications. In *2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC)* (pp. 1–8). IEEE.

[46] Bonner, E., & Reinders, H. (2018). Augmented and virtual reality in the language classroom: Practical ideas. *Teaching English with Technology, 18*(3), 33–53.

[47] Nguyen, Q. K., & Dang, Q. V. (2018). Blockchain technology for the advancement of the future. In *2018 4th International Conference on Green Technology and Sustainable Development (GTSD)* (pp. 483–486). IEEE.

[48] Alam, M. F., Katsikas, S., Beltramello, O., & Hadjiefthymiades, S. (2017). Augmented and virtual reality based monitoring and safety system: A prototype IoT platform. *Journal of Network and Computer Applications, 89*, 109–119.

[49] Khor, W. S., Baker, B., Amin, K., Chan, A., Patel, K., & Wong, J. (2016). Augmented and virtual reality in surgery—the digital surgical environment: Applications, limitations and legal pitfalls. *Annals of Translational Medicine, 4*(23), 454.

[50] Abu-Zaid, A., Aljaidi, M., Al-Na'amneh, Q., Samara, G., Alsarhan, A., & Qadoumi, B. (2025). Advancements and challenges in the internet of drones security issues: A comprehensive review. In *Machine Intelligence Applications in Cyber-Risk Management* (pp. 1–24). IGI Global.

[51] Al-Na'amneh, Q., Aljaidi, M., Gharaibeh, H., Nasayreh, A., Al Mamlook, R. E., Almatarneh, S., Alzu'bi, D., & Husien, A. S. (2023). Feature selection for robust spoofing detection: A chi-square-based machine learning approach. In *2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)* (pp. 1–7). IEEE.

[52] Al-Na'amneh, Q., Nasayreh, A. N., Al Mamlook, R., Gharaibeh, H., Alsheyab, A. M., & Almaiah, M. (2024). Improving memory malware detection in machine learning with random forest-based feature selection. In *Risk Assessment and Countermeasures for Cybersecurity* (pp. 96–114). IGI Global.

[53] Alghamdi, A. (2024). Exploring the security landscape of AR/VR applications: A multi-dimensional perspective.

[54] Ali, S., Alshinwan, M., Khashan, O. A., Hijjawi, M., Altawil, A., Al-Na'amneh, Q., Abu-Adaiq, H., Alhardan, H., AbdElminaam, D. S., Tarawneh, O., et al. (2025). Intrusion detection for wireless sensor networks using parrot algorithm. In *Machine Intelligence Applications in Cyber-Risk Management* (pp. 345–366). IGI Global Scientific Publishing.

[55] Alismail, A., Altulaihan, E., Rahman, M. H., & Sufian, A. (2022). A systematic literature review on cybersecurity threats of virtual reality (VR) and augmented reality (AR). In *Data Intelligence and Cognitive Informatics: Proceedings of ICDICI 2022* (pp. 761–774).

[56] Alismail, A., Altulaihan, E., Rahman, M. H., & Sufian, A. A systematic literature review MKKS on cybersecurity threats of virtual reality (VR) and augmented reality.

[57] Kurii, Y., & Opirskyy, I. (2022). Analysis and comparison of the NIST SP 800-53 and ISO/IEC 27001: 2013. *NIST Special Publication, 800*(53), 10.

Biographies

Dr. Mohammed Amin is among the top 2% scientists in the world from 2020 up to now. He is working as Editor in Chief for the International Journal of Cybersecurity and Risk Assessment. He has published over 115 research papers in highly reputed journals such as the Engineering and Science Technology, an International Journal, Education and Information Technologies, IEEE Access and others. Most of his publications were indexed under the ISI Web of Science and Scopus. His current research interests include Cybersecurity, Cybersecurity-Risk Assessment and Blockchain. m.almaayah@ju.edu.jo
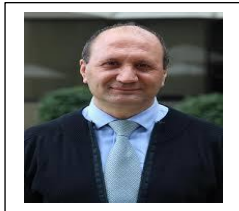
Prof. Youakim Badr received his bachelor's and master's degrees in computer science from the Lebanese University and an additional master's degree in mathematical modeling and scientific software engineering from the Francophone University Agency (AUF). He earned his Ph.D. in computer science from the National Institute of Applied Sciences (INSA-Lyon), where he worked as an associate professor in the computer science and engineering department. https://orcid.org/0000-0002-8976-7894

Dr. Mahmoud Mohammad Aljawarneh. Received the B.S., M.S. and PhD degree in Information Technology from University of Sindh, Jamshoro, Pakistan, in 2011, 2014, and 2019 respectively. From 2012 to 2014 and 2016 to 2018, he was a Research Assistant with the Institute of Information and Communication Technology, University of Sindh, Jamshoro, Pakistan. And From January 2020 to March 2022: he was an Assistant Professor, Computer Science Department, Szabist University, Hyd Campus, Pakistan. And From October 2020 to Current: he is working as an Assistant Professor, Faculty of Information Technology, Applied Sciences Private University, Amman, Jordan. His research interest includes Pervasive Computing, Context-awareness, Internet of Things (IoT), Machine Learning (ML). ma_jawarneh@asu.edu.jo
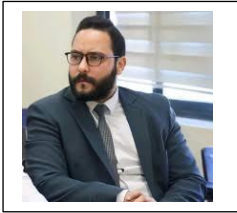
Qais Al-Na'amneh received his B.S. in software engineering from Hashemite university, Jordan, in 2015; the M.Sc. (with distinction) in Cyber Security from Hashemite university, Jordan, in 2021; I. Al-Naamneh is currently working as an Instructor with the Cyber Security and Cloud Computing Department, Applied Science Private University, Jordan. Also I have many certificates in Information Technology field like MCSA 2012 (Microsoft Certified Solutions Associate),MS-500 (Microsoft 365 Security Administration Certificate),CCNA (Cisco Certified Network Associate),CSFPC (Cyber Security Foundation Professional Certificate),Yeastar Certified Technician (VoIP phone system),Oracle developer, CEH (Ethical Hacker). q_naamneh@asu.edu.jo
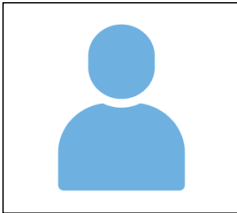
Rahaf Hazaymih received her B.Sc. in Computer Engineering from Jordan University of Science and Technology in 2020. She is currently pursuing a Master's degree in Data Science. Rahaf is working as an AI Engineer, with research interests in Natural Language Processing (NLP), Artificial Intelligence (AI), Machine Learning (ML), Data Science, Cybersecurity, and the Internet of Things (IoT). Rahaf_hazaymih@yahoo.com
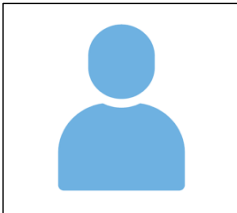
Shahid Munir Shah, Faculty of Engineering Sciences and Technology, Hamdard University, Karachi, Pakistan. shahid.munir@hamdard.edu.pk