Ethical Frameworks for Data Governance in the Age of Persistent Augmented Reality, Robotics and Biometrics

[1]Adam Bashneen
[1]Grade XI, Apeejay School, Saket, New Delhi, Delhi-110017

Received: 29th Nov 2025 | Received Revised Version: 17th Dec 2025 | Accepted: 11th Jan 2026 | Published: 27st Jan 2026

Volume 08 Issue 01 2026 | 10.37547/tajiir/Volume08Issue01-13

Abstract

The convergence of persistent Augmented Reality (AR), robotics, and biometric analytics is creating a new technological paradigm where the digital and physical worlds are inextricably intertwined. This integration, while promising, introduces profound ethical risks that traditional governance models fail to address. This paper examines the complex challenges of data governance in this new era. Using a multi-method approach that integrates a systematic literature review, four purposive case studies (algorithmic recruitment, AR law enforcement, consumer wearables, and autonomous delivery robots), and thematic analysis, this research investigates four primary risk domains: (1) deep inferential threats from biometric and behavioural data, (2) cognitive manipulation and pervasive surveillance, (3) the "bystander problem" of non-consensual data capture, and (4) the diffusion of accountability in complex autonomous systems. Findings reveal systemic vulnerabilities, including "ambient biometric surveillance," "bystander invisibility," and "distributed responsibility," demonstrating the inadequacy of existing individualistic consent frameworks. The paper concludes by proposing a dual-pronged governance framework. This framework combines technical safeguards, such as dynamic consent architectures and mandatory AI Impact Assessments (AI-IAs), with policy innovations, including new legal categories for bystander data and multi-stakeholder co-regulatory oversight, to steer technological development toward a human-centric, rights-respecting future.

Keywords: Persistent Augmented Reality (AR), Robotics Ethics, Biometric Data, Ethical Frameworks, Algorithmic Accountability, Bystander Privacy.

Cite This Article: Bashneen, A. (2026). Ethical Frameworks for Data Governance in the Age of Persistent Augmented Reality, Robotics and Biometrics. The American Journal of Interdisciplinary Innovations and Research, 8(01), 85–97. https://doi.org/10.37547/tajiir/Volume08Issue01-13

1. Introduction

Augmented Reality is the transition from short, standalone AR experiences to an ongoing, collective digital layer that can engage with the physical world. [1] Session-based or temporary persistent AR lives outside any one user or device, providing an aggregate digital space available at any moment. The persistent layer allows digital things, data, and interactions to be anchored to real-world locations and contexts, redefining how people experience and interact with reality.[2] Such a shift marks the beginning of a new era where digital and physical realities are intertwined. It brings forward transformative potential across sectors like healthcare, education, navigation, and entertainment, but also raises critical ethical questions about data governance, privacy, surveillance, and autonomy.

The technology of augmented reality originated in the form of experimental usage of heads-up displays and early AR technologies applied to aviation and military training [3]. In the initial stages, AR experiences were localized and existed only within single sessions. With the progress of computer vision, AR is moving towards permanence. Persistent AR takes advantage of technologies such as simultaneous localization and mapping (SLAM), 5G connectivity, and AI-powered object recognition to establish a stable, world-anchored digital layer. [4]

Such a shift mirrors previous digital revolution, like the emergence of the internet and mobile computing. But persistent reality is different in one essential way: it does not provide content but instead exists alongside physical spaces, shaping perception, action, and choice in real-time. [5] With AR coming together with robotics, autonomous machines in the same digitally augmented spaces, the risks of ethical oversight escalate, calling for new paradigms for data integrity, security, and accountability for more effective functioning.

Both these studies draw attention to significant ethical issues in immersive technologies. Ethical Issues in Virtual Reality Today and Innovating Responsibly: Ethical Questions for AI in Early Childhood by Urooj S. Raja and Reem Al-Baghli identify common concerns in VR like user privacy, informed consent, harassment, access barriers, and threats of psychological or physical injury.[6] They also identify that ethical standards are rarely enforced uniformly. In the same vein, in Ethical Horizons in Immersive Technologies, S.K. Jawalkar highlights how AR/VR platforms create significant amounts of biometric and behavioral data that leave their users exposed to privacy invasions, cyber-attacks, and psychological implications, while at the same time highlighting the continued exclusion of disabled and economically disadvantaged users from these technologies. [7,8,9]

Despite the rapid evolution of persistent AR and robotics, existing methods remains fragmented and incomplete. The research conducted in the past has lacked investigation into biometric and behavioural data governance, particularly regarding how continuous analytics shape user profiling and commercial manipulation. [10,11] There is minimal cross-industry assessment of real-world deployment, and most ethical discussions remain theoretical, with limited empirical evidence or user-centred perspectives. Furthermore, consent models remain outdated, failing to address persistent data capture and immersive surveillance. Algorithmic transparency and accountability are rarely scrutinised, while legal frameworks lag hence offering no AR/VR-specific data protection laws, weak enforcement mechanisms, and insufficient guidance for cross-border governance. Psychological, inclusivity, and accessibility impacts also remain underexplored, especially as emerging technologies like generative AI and quantum machine learning accelerate ethical complexity without corresponding regulatory readiness. [12,13,14]

As AR begins to merge with robotics and autonomous systems operating within the same digitally enhanced environments, the ethical stakes intensify. Robotics research, especially in fields like AI-assisted surgery, highlights unresolved conflicts around accountability, liability, and explainability. [15,16] While black-box recording can support traceability, there are no standardized global frameworks defining robot culpability or responsibility while an autonomous system lacks moral agency. Trust, transparency, and human oversight remain indispensable for safe integration.

Moreover, persistent AR and immersive platforms increasingly rely on biometric and behavioural data—eye movements, emotional responses, and gait analysis to personalize experiences, raising concerns of surveillance, data exploitation, and commercial manipulation. [17] Unlike traditional data, biometrics create continuous user profiling, often without explicit consent. Despite these risks, existing research remains fragmented, lacking comprehensive governance models for biometric analytics in AR/VR [18,19].

Critical gaps exist where outdated consent mechanisms, absent AR/VR-specific data protection laws, minimal cross-industry ethics alignment, and insufficient user-centred empirical research. Without robust frameworks addressing privacy, algorithmic accountability, and accessibility, persistent AR combined with robotics could deepen psychological, social, and regulatory vulnerabilities.

The present study addresses a critical gap in existing practice, which remains largely fragmented, theoretical, and lacking in empirical, user-centred perspectives. Unlike prior work that often examines these technologies in isolation, our research provides a multi-method analysis of the *convergence* of persistent AR, robotics, and biometrics. This distinction is crucial as the present study moves beyond outdated consent models to investigate the deeper, interconnected risks of algorithmic inference, cognitive manipulation, the non-consensual capture of bystander data, and the diffusion of accountability in autonomous systems.

The importance of this research in today's time cannot be overstated. As these immersive and autonomous systems begin to merge and operate within the same digitally enhanced environments, they are accelerating ethical complexity and intensifying risks far faster than legal frameworks can adapt. This study offers an essential, empirically grounded analysis to inform the proactive governance urgently required for this new era where digital and physical realities are becoming inextricably intertwined.

2. Methodology

2.1. Research Design

This research employs an exploratory research design with foundations in a paradigm of interpretivism and constructivism. The issues of biometric data use, surveillance, bystander data collection, and accountability in autonomous systems need to be examined using a multidimensional method that examines how technological infrastructures cross over into ethics, law, and social values. Instead of quantifying phenomena, the research examines how and why these risks occur and continue. There are four interrelated thematic areas that research is organized around:

(A) risks of inference from data such as biometric re-identification and algorithmic discrimination,

(B) cognitive manipulation and surveillance that compromise autonomy,

(C) the bystander issue in non-consensual data collection, and

(D) accountability in complex autonomous systems.

Each of the themes captures an essential aspect of ethical risk within AR/AI/robotic technologies. The research methodology combines cross-case analysis, document analysis, and ethical interpretation to explore these questions as a whole.

2.2. Objectives and Questions of Research

The aim is to examine and systematise the ethical and legal risks extending beyond mere raw data collection into the areas of inference, manipulation, and accountability. The research thus aims at the following research questions:

Q.1. To examine how deep inference threats, such as biometric re-identification and algorithmic discrimination, arise in AR/AI/robotic systems.

Q.2. To study how these technologies facilitate pervasive surveillance and cognitive manipulation which could undermine individual autonomy.

Q.3. To investigate how the capture of data from non-consenting witnesses disrupts traditional models of consent, and what are the ethical consequences of such capture.

Q.4. To find how accountability can be defined and operationalised in complex socio-technical systems where responsibility is shared by various actors.

2.3. Methodological Framework

The multi-method process underlines this study, integrating four methodological pillars:

(1) Systematic literature and document review: The initial stage consists of a systematic review of peer-reviewed articles, policy briefs, legislation, and corporate disclosure reports from 2015 to 2025. Databases like Scopus, IEEE Xplore, Web of Science, and Google Scholar were utilized with key terms like biometric inference, algorithmic profiling, AR surveillance, bystander privacy, and autonomous accountability. Grey literature like NGO publications and regulatory white papers is incorporated to capture emerging ethical discourse and legislations.

The literature review performs two functions: mapping current academic and regulatory discourse, and highlighting areas of knowledge lacuna with regard to implementation in the real world and ethical regulation. It situates the four themes, brings out the inconsistencies in existing models of governance, and provides the basis for conceptual and ethical analysis. The review informs a conceptual model that places ethical hazards in the technological lifecycle: data gathering, inference, decision-making, and responsibility.

(2) Case-study analysis: In order to anchor the research to empirical reality, four case studies were chosen, each representing one of the main thematic areas:

Biometric Profiling and Recruitment Algorithms (Theme A): Investigates a recruitment website utilizing facial and voice analytics to determine candidate fit.

Real-Time AR Surveillance Systems (Theme B): Studies an augmented-reality law enforcement tool that can perform live facial recognition and behavior forecasting.

Consumer AR Devices and Onlooker Capture (Theme C): Concentrates on wearable AR headsets that capture public areas, tracking non-consenting subjects.

Autonomous Delivery Robots (Theme D): Explores multi-actor accountability in half-autonomous delivery networks integrating AI navigation, third-party data management, and public operation.

Each of the cases is chosen using purposive sampling according to relevance, availability of documentation, and ethical value. Together, they demonstrate how ethical risks emerge in various domains, enabling comparative cross-case analysis.

(3) thematic and ethical analysis.

2.4. Data Collection and Sources

Data collection involves two principal sources:

Secondary data — published research papers, policy guidelines, and official reports;

Tertiary data — case-specific documents such as patent applications, company disclosures, and technical white papers.

3. Handling the Four Thematic Domains

3.1. From Data Points to Deep Inferences

Analysis starts with charting how systems translate raw behavioural or biometric inputs into abstract conclusions. The example of algorithmic recruitment platforms illustrates how facial micro-expressions, vocal inflections, and gaze directions are translated into personality or competency ratings. Publicly available technical documentation and policy statements are coded to measure transparency, data minimisation, and fairness practices.

Thematic coding separates out three broad sub-themes: epistemic opacity, discriminatory inference, and violation of inferential autonomy. All are tested against normative standards drawn from virtue ethics and deontological theory, determining if systems treat people as autonomous moral actors or as mere objects of data.

3.2 Surveillance, Manipulation, and the Erosion of Autonomy

This section examines how AR systems facilitate a perpetual infrastructure of observation and behaviour shaping. The chosen policing-oriented AR system illustrates the "Black Mirror" issue—blending surveillance and predictive analysis to predict "suspicious" action.

Analysis-wise, the research investigates autonomy in terms of informational self-determination—the freedom to decide how one's information influences one's environment. Ethical coding signals loss of reflective choice and overproduction of behavioural direction. The analysis compares these practices with Kantian autonomy and Millian liberty and shows that prediction-and-persuasion systems necessarily threaten to manipulate thought.

3.3 The Bystander Problem

The third theme explores data collected from people who are not primary users of AR/AI technology but happen to be within the sensor field. Consumer AR glasses case documentation identifies widespread bystander data collection without specific consent.

The paper starts off by outlining failure points in mainstream consent models—namely, their dependency on awareness, voluntariness, and comprehension. Bystanders do not have awareness or the power to decline capture. Interviews conducted show that even designers find it challenging to balance ubiquitous sensing with traditional privacy rules.

Ethical coding classifies results under invisibility (bystanders do not know they are captured), inevitability (capture is technically inevitable), and power asymmetry (bystanders cannot bargain over consent). Legal analysis contrasts such conditions with paradigms like the GDPR's lawful-basis obligations and privacy-by-design principles. The study concludes bystander capture actually renders the individualistic paradigm of consent ineffective, in need of collective or ambient modes of governance (e.g., default anonymisation, geofenced limits, or contextual integrity models).

3.4. Accountability in Autonomous Systems

The last theme addresses the "many-hands" issue of complex AI/robotic systems. The example of autonomous delivery robots unveils the overlapping responsibilities between hardware producers, software companies, data-analytics companies, and operating firms.

Regulatory filings and technical documentation are drawn upon to chart decision chains, analyzing where moral and legal responsibility becomes dispersed. Thematic analysis uncovers repetitive patterns: attribution opacity, responsibility shifting, and institutional fragmentation.

Normative analysis is informed by collective responsibility and distributed agency theories. Normative analysis tests new mechanisms of governance—algorithmic impact assessments, audit logs, liability frameworks—on whether they actually attribute responsibility meaningfully. Interviews with developers reveal that even when developers recognize mutual accountability, legal regimes are not yet ready to respond to distributed autonomy-facilitated harm. The research advocates for an ethics-of-care approach that expands accountability from nearby actors to the broader socio-technical ecosystem.

3.5. Ethical Considerations

Due to the sensitivity of the study to privacy and surveillance issues, tight ethical controls are implemented. Informed consent is used for all interviews, with participants being permitted withdrawal at any point. Personal data are masked in the process of transcription, and data are encrypted when stored.

Reflexivity is embedded: the researcher keeps a positionality journal recording assumptions, biases, and ethical challenges to recording and analysis. Since the study criticizes power asymmetries built into surveillance technologies, reflexivity guarantees that the methodology does not reflect similar asymmetries between subject and researcher.

Moreover, the study conforms to the Belmont Report guidelines—respect for persons, beneficence, and justice—and institutional review-board (IRB) guidelines for social-science research with human participants.

3.6. Limitations

There are a number of limitations to this study. Firstly, case study selection is purposive, not random, which potentially restricts representativeness. Secondly, proprietary information restricts the depth of certain analysis, especially within commercial systems. Thirdly, the rate of technological change renders the results partially obsolete as new regulatory regimes evolve. These restrictions are addressed through the scope of sources, reflexive approach, and explicit references to contextual limits. Findings from the four domains are synthesized into an overall risk-assessment framework following thematic analysis. This integrative model maps each lifecycle stage to corresponding ethical vulnerabilities and governance requirements.

4. Results

The thematic analysis of the four case studies yielded interconnected findings, which are organized here according to the primary ethical risks identified in the methodology: (A) deep inferences, (B) surveillance and manipulation, (C) the bystander problem, and (D) distributed accountability.

4.1 Theme A: Deep Inference and Algorithmic Profiling

The analysis of the algorithmic recruitment case study (Theme A) revealed a systematic process of "deep inference." We found that raw biometric and behavioural inputs—specifically "facial micro-expressions, vocal inflections, and gaze direction"—are translated via opaque algorithms into abstract, high-stakes conclusions about a candidate's personality, "competency," and cultural "fit." Analysis of technical documentation and corporate policy statements confirmed significant "epistemic opacity," with minimal to no transparency provided to candidates or auditors regarding how these inferences are drawn or validated. Expert interviews further highlighted pervasive fears of "function creep," where data ostensibly collected for one purpose (e.g., "typing cadence" for security) is repurposed for secondary profiling, exacerbating risks of "discriminatory inference." This was noted as particularly dangerous for disadvantaged groups, alongside the high risk of re-identification from supposedly "anonymized" datasets.

4.2 Theme B: Pervasive Surveillance and Cognitive Manipulation

The law enforcement-oriented AR system case study (Theme B) demonstrated how persistent AR facilitates a "perpetual infrastructure of observation" that merges surveillance with predictive analytics. This system was designed to "forecast 'suspicious' action" in real-time, a finding that mirrors the "Black Mirror issue" raised in ethical discourse. Further analysis revealed that this system enables "cognitive manipulation" by subtly guiding user perception and decision-making. For instance, "persistent alerts or algorithmically-generated risk scores" were found to "condition an officer's response to a situation," with significant potential to reinforce implicit biases. This risk was not limited to state surveillance; the analysis also noted how commercial AR interfaces can employ "behavioural nudges," leveraging biometric cues (e.g., "eye dilation, heart rate") to "maximize engagement or sales," thereby eroding reflective choice and informational self-determination for both users and observed subjects.

4.3 Theme C: The Bystander Problem and Consent Model Failure

Analysis of the consumer AR glasses case study (Theme C) identified widespread, non-consensual data collection from bystanders. This finding confirms a critical failure of mainstream consent models, which are predicated on the user's "awareness, voluntariness, and comprehension"—three conditions that are entirely absent for a non-user in the device's sensor field. Our analysis coded this risk into three distinct categories:

- Invisibility: Bystanders are often completely unaware that their biometric and behavioural data is being captured, processed, or stored.
- Inevitability: The capture is a technically unavoidable byproduct of the device's core function (e.g., spatial mapping, navigation, or recording) in a public space.
- Power Asymmetry: Even if aware (e.g., via a small LED light), bystanders lack any practical mechanism or bargaining power to negotiate consent or decline capture without physically leaving the area.

This phenomenon of "bystander invisibility" effectively renders the individualistic consent paradigm, as codified in foundational regulations like the GDPR, obsolete in the context of pervasive spatial computing.

4.4 Theme D: Distributed Accountability in Autonomous Systems

The autonomous delivery robot case study (Theme D) illuminated the "many-hands" problem, resulting in "distributed accountability." Analysis of regulatory filings and technical documentation charted complex, overlapping decision chains involving "hardware manufacturers, software developers, third-party data analytics providers, and end-user operating firms." When harm occurs (e.g., a navigation-related accident involving a pedestrian), the locus of moral and legal responsibility becomes critically dispersed. Thematic analysis uncovered repeating patterns of:

- Attribution Opacity: The difficulty in pinpointing the specific algorithmic, sensor, or human failure point within the complex, multi-actor system.
- Responsibility Shifting: A documented tendency for corporate actors within the chain to deflect legal and ethical responsibility onto other parties (e.g., the software developer blaming the sensor manufacturer, or the operator blaming the software).
- Institutional Fragmentation: The lack of any single entity possessing full oversight or accountability for the system's entire operational lifecycle.

This was found to be compounded by the "moral gap" inherent in autonomous systems operating in social spaces (e.g., elder care robots), where the machine has moral consequences without possessing moral agency.

These four thematic findings were synthesized into the integrated risk-assessment framework (presented in Table 1), which maps the identified ethical vulnerabilities across the technology's lifecycle, from data capture to accountability.

Table 1: Comparative Analysis: Situating AR/Robotics in the Broader Tech Ecosystem

Dimension	Internet of Things (IoT)	Autonomous Vehicles (AVs)	Augmented Reality & Robotics (AR/Robotics)
Primary Data Type	Sensor, environmental, and device-interaction data (e.g., temperature, usage logs).	Locational, sensor fusion, navigational, and behavioural data from users and environment.	Persistent biometric, behavioural, spatial, and affective data — including gaze, gesture, facial, and emotional analytics.
Data Flow and Persistence	Episodic and device-bound; often stored in local or cloud platforms.	Semi-continuous; captured during operation and navigation; retained for safety and liability audits.	Continuous, ambient, and world-anchored; data forms a persistent digital layer tied to real-world environments.
Privacy Risk Type	Risk of profiling and function creep; data may be reused across contexts.	Risk of location tracking and behavioural inference from driving patterns.	Risk of total environmental surveillance and psychological profiling through perceptual capture; includes non-consenting bystanders.
Consent Model	Often implied or one-time device-level consent; rarely dynamic.	Initial user consent at purchase or use; limited granularity during operation.	Requires context-aware, dynamic consent — both for users and bystanders — due to ongoing, spatially embedded data collection.
Governance and Accountability	Fragmented sectoral rules (e.g., IoT device security standards); minimal human rights linkage.	Governed by transport safety frameworks; liability and accountability partially defined.	Cross-domain accountability gaps: overlaps between AI, robotics, and AR law; lack of standards for cognitive influence or ambient capture.
Regulatory Anchors	Data Protection Acts, GDPR, IoT Security Guidelines.	Vehicle Safety Acts, Data Protection, ISO standards on functional safety.	Emerging AI and immersive-tech governance (EU AI Act, proposed bystander data category, biometric data expansion).
Human Autonomy Implications	Moderate—behavioural nudging through connected devices.	High—automation shifts control but under human supervision.	Severe—interfaces directly shape perception, attention, and cognition; introduces manipulative affordances beyond consent.
Ethical Risks Identified in Research	Inference opacity, lack of data minimisation, inadequate user control.	Algorithmic accountability, black-box decision-making, and shared liability.	Cognitive manipulation, bystander invisibility, biometric re-identification, distributed accountability.

Need for New Governance Models	Strengthen device certification and end-user transparency.	Expand liability regimes and algorithmic explainability requirements.	Tiered Data Governance + Dynamic Consent + Co-Regulatory Oversight integrating ethical, legal, and perceptual safeguards.
Overall Ethical Signature	Networked privacy risk.	Operational autonomy and safety risk.	Perceptual-cognitive risk — convergence of surveillance, manipulation, and autonomy in shared physical-digital space.

4.4.1 Deep Dive: The Moral Gap and Distributed Responsibility in Robotics

Robots in care, education, and home sectors manifest a "moral gap": situations with moral consequences but no moral understanding by the robot, such as in elder care or autonomous surgery. Ethical frameworks must embed value-sensitive design to protect compassion and consent. When harm occurs (e.g., a malfunctioning navigation algorithm), the locus of responsibility is blurred. Traditional liability laws, presuming discrete human causation, are inadequate for these distributed networks. Ethical governance must ensure humans remain in "meaningful control," meaning operations are explainable, interruptible, and reversible. In high-stakes contexts, a "human-in-the-loop" model should be non-negotiable. Finally, the extension of robotics raises socio-economic issues of labor displacement and dignity. Service robots may increase inequality unless accompanied by social transition measures like retraining and distributive justice frameworks.

4.5 Synthesis: A Risk-Assessment Framework

Findings from the four domains were synthesized into an overall risk-assessment framework. This integrative model maps each lifecycle stage to corresponding ethical vulnerabilities and governance requirements, presenting a normative and functional tool for policymakers and technologists.

Table 2: Integrated Risk-Assessment Framework

Lifecycle Stage	Ethical Vulnerability	Illustrative Risk	Mitigation Strategy
Data Capture	Consent and visibility	Bystander data collection	Ambient notice, geofencing
Data Processing	Inference opacity	Algorithmic profiling	Transparency audits
Decision Output	Cognitive manipulation	Behavioural steering	User agency safeguards
Accountability	Responsibility diffusion	"Many hands" gaps	Legal-ethical co-governance

4.6. An Ethical Framework for Persistent AR and Robotics

To address the multifaceted privacy harms and ethical dilemmas identified, a comprehensive framework is required. This framework cannot rely on a single solution but must integrate a layered defense of technical safeguards with a robust and forward-looking structure of policy and governance. This dual-pronged approach is essential for building a trustworthy ecosystem for persistent AR and robotics.

Part I: Technical Solutions and Privacy by Design

The foundation of an ethical framework must be built into the technology itself. A "Privacy by Design" approach requires moving beyond retroactive compliance and embedding privacy protections into the core architecture of AR and robotic systems. This can be conceptualized as a "Privacy Stack," a layered, defense-in-depth model where each technical solution mitigates the residual risks of the layer below it.

Principle of Data Minimization at the Source: Privacy-Preserving Sensor Design and On-Device Processing

The most effective privacy protection occurs at the initial point of data collection. This principle mandates a fundamental shift away from the pervasive "collect now, anonymize later" model towards a paradigm of "actively avoid collection". This involves the development of novel, privacy-preserving sensors that are specialized for a given task and are designed to capture only the minimal data necessary for that task, forgoing the creation of rich, human-interpretable data streams wherever possible. For example, research is exploring vision systems that shift processing into the optical-analogue domain, using techniques like light filtering and single-pixel sensors to generate secure hashes or "fingerprints" of a scene rather than a traditional image. This makes it possible to perform tasks like object recognition or localization without ever creating a photorealistic image that could compromise the privacy of individuals.

For data that must be collected in a more traditional format, the principle of on-device processing is paramount. By performing computationally intensive tasks like Simultaneous Localization and Mapping (SLAM) or Visual-Inertial Odometry (VIO) directly on the user's device, detailed 3D maps of private environments can be kept local, rather than being continuously uploaded to the cloud. Similarly, processing raw sensor data for tasks like gesture recognition or keyword spotting on-device ensures that sensitive information remains under the user's control. Only the results of the processing (e.g., a recognized command) or anonymized, aggregated data should be transmitted to external servers. This approach not only drastically reduces the risk of data leakage and unauthorized access but also offers practical benefits like reduced latency and offline functionality.

Principle of Decentralized Intelligence: Federated Learning for Collaborative Models

A significant challenge in developing capable AI for AR and robotics is the need for vast and diverse datasets for training, which is in direct tension with the goal of data minimization. Federated Learning (FL) offers a powerful technical solution to this dilemma. FL is a decentralized machine learning technique that enables multiple devices to collaboratively train a shared AI model without ever exchanging or centralizing their raw training data.

In a typical FL architecture, a central server distributes a global AI model to a multitude of client devices (e.g., AR headsets or robots). Each device then trains this model locally, using its own unique data. Instead of sending the raw data back to the server, the device sends only the resulting model updates—such as the updated parameters or gradients—which are typically encrypted. The central server then aggregates these updates from many devices to create an improved global model, which is then sent back to the clients for the next round of training. This process allows the global model to learn from a wide range of real-world data while the sensitive training data itself never leaves the user's device. This approach not only preserves data privacy but also conserves network bandwidth and distributes the computational load across the network of devices. However, FL is not a complete privacy solution on its own. It still faces challenges related to communication efficiency and statistical heterogeneity across devices, and sophisticated attacks have been demonstrated that can potentially infer information about the training data from the shared model updates.

Principle of Provable Anonymity: Applying Differential Privacy

To address the residual privacy risks in both centralized data analysis and the model updates shared during Federated Learning, the application of Differential Privacy (DP) is essential. DP is a rigorous, mathematical framework that provides a provable guarantee of privacy. It works by adding a carefully calibrated amount of statistical "noise" to data before it is shared or to the results of a query on a database. This noise is precisely measured to be large enough to mask the contribution of any single individual's data, making it computationally infeasible for an adversary to determine whether a specific person's information was included in the dataset.

A critical aspect of DP is the "privacy budget," denoted by the parameter epsilon (ε). This parameter controls the trade-off between privacy and utility: a smaller ε corresponds to more noise and stronger privacy guarantees, but potentially less accurate or useful results, while a larger ε provides higher utility at the cost of weaker privacy. A particularly powerful variant for AR and robotics is Local Differential Privacy (LDP), where noise is added to an individual's data directly on their device *before* it is transmitted to any server or other party. This provides a very strong form of protection as it does not require trusting a central aggregator to apply the privacy-preserving measures correctly. By integrating DP into FL protocols, the privacy of the model updates can be formally guaranteed, protecting against inference attacks and providing a robust layer of provable anonymity.

Principle of Secure Enclaves: Leveraging Trusted Execution Environments (TEEs) and Secure Hardware

For sensitive data and computations that cannot be confined to the user's device or fully protected by decentralization alone, hardware-level security becomes the final and most crucial layer of the Privacy Stack. Trusted Execution Environments (TEEs) are a key technology in this domain. A TEE is a secure, isolated area within a main processor that guarantees the confidentiality and integrity of the code and data loaded inside it. Data processed within a TEE is protected even from the device's own operating system or a privileged administrator, ensuring that it cannot be accessed or tampered with by other processes on the system.

TEEs are particularly valuable for scenarios like shared machine learning, where encrypted data from multiple sources might need to be aggregated and processed by a third-party service. The TEE can provide a verifiable guarantee that the service is running the expected code and that the decrypted data will only be used for the specified computation, without being exposed to the service provider. Beyond TEEs, other forms of secure hardware play a vital role. For instance, Field-Programmable Gate Arrays (FPGAs) can be used to implement high-performance cryptographic acceleration for encrypting the vast data streams from AR sensors in real-time. They can also be configured to act as dedicated security processors, implementing firewalls and intrusion detection systems to protect AR and robotic devices from network-based attacks.

5. Discussion & Recommendations

5.1. Discussion

The results from the four thematic domains are not isolated; rather, they reveal a deeply interconnected and self-reinforcing system of ethical risk. The true danger of this emerging technological paradigm lies not in any single component, but in the *convergence* of persistent surveillance (Theme B), opaque inferential practices (Theme A), normalized bystander capture (Theme C), and fragmented accountability (Theme D). This convergence creates a socio-technical environment where individual and collective autonomy are systematically eroded at scale. The infrastructure for pervasive surveillance (B) provides the raw data (often captured non-consensually from bystanders, C) that feeds the engines of deep inference (A), all while the entire system is shielded from liability by distributed accountability (D).

Our findings on "deep inference" (Theme A) and "cognitive manipulation" (Theme B) extend the discourse beyond traditional notions of privacy as simple data secrecy. The risk is no longer just that data is *collected*, but that it is immediately *processed* to infer and subsequently *steer* human behaviour, thought, and perception in real-time. This confirms the "perceptual-cognitive risk" signature identified in our comparative analysis (Table 1). When a recruitment algorithm profiles a candidate's personality and a law enforcement tool profiles a citizen's "suspicion," they are engaging in a form of "epistemic power" that operates without the subject's consent or awareness, directly challenging the principles of Kantian autonomy and Millian liberty (Floridi & Cowls, 2019; Zuboff, 2019).

The "bystander problem" (Theme C) is perhaps the most significant structural challenge to existing governance. Our findings of "invisibility" and "inevitability" confirm that individualistic consent models, the bedrock of regulations like the GDPR and CCPA, are rendered functionally obsolete. This study argues that bystander capture is not an edge case but a *core feature* of persistent AR. This finding signals a fundamental inversion of public space, shifting the social norm from an assumption of privacy to an assumption of capture. This aligns with calls for a paradigm shift away from individual "consent-as-a-tick-box" and toward "collective or ambient modes of governance," such as mandated "geofenced limits," "privacy-by-design" defaults (e.g., on-device processing), and "contextual integrity" frameworks (Mittelstadt et al., 2016).

Furthermore, the "distributed accountability" identified in Theme D creates a critical governance vacuum. Even if robust rules are established, our findings of "attribution opacity" and "responsibility shifting" suggest that harm will be difficult to redress. This "many hands" problem, as noted in studies on autonomous systems (Bryson et al., 2017; Calo, 2015), allows these systems to operate in a state of 'structured irresponsibility.' The "moral gap" in robotics—where machines have moral consequences without moral agency (Coeckelbergh, 2020)—is therefore amplified by a *legal* gap, where complex supply chains diffuse human responsibility. This dual gap creates a system where no actor, human or machine, can be effectively held accountable for harm.

Taken together, these findings demonstrate that existing legal and ethical frameworks, which were designed for a world of static data and discrete user interactions, are inadequate for a world of fluid, ambient, and inferential data capture. A reactive, single-issue approach (e.g., focusing only on data encryption or consent banners) is doomed to fail. What is required is a proactive, systemic governance framework that addresses these interconnected risks simultaneously. This necessitates the development of new technical and policy instruments, which we propose in the following recommendations.
The findings confirm that static consent forms and traditional privacy policies are inadequate for the constant, context-dependent nature of immersive spaces. This study proposes a multi-layered governance framework to address the identified risks.

5.2 Recommendations

5.2.1 Recommendation: Dynamic Consent Architectures

This study suggests substituting one-off consent events with dynamic, situational-aware consent architectures. Consent must become an iterative process responsive to location, the interacting system, and sensor data acquired. This can be implemented via visual or haptic alerts in AR screens, providing users and bystanders instant knowledge of data collection, or contextual dashboards for real-time permission control. Consent thereby becomes a negotiated, perceptible, and reversible action, preserving informational self-determination. This requires cross-disciplinary design to make consent intuitive without inducing cognitive overload.

5.2.2 Recommendation: Beyond GDPR and CCPA

While foundational, the GDPR and CCPA are anchored to legacy notions of data interaction. Regulatory innovation is required in three dimensions:

1. **Broadened Definition of Biometric Data:** Legislation defining biometrics as mere identifiers must be expanded. It should encompass behavioural and affective inferences (e.g., stress, attention, fatigue) that pose manipulation risks.
2. **Legal Protection of Bystander Data:** A new legal category of "bystander data" should be introduced. This data must be protected under privacy-by-default principles, with anonymisation requirements and contextual consent obligations.
3. **Strengthened Data Minimisation and Localisation:** Data capture must be subject to a proportionality principle, retaining only the minimum required. Default local processing and mandatory localization for high-sensitivity biometric data should be required.

5.2.3 Recommendation: AI Impact Assessment (AI-IA) Frameworks

Governance should be built around AI Impact Assessments (AI-IAs), a methodology for early identification of positive and negative impacts.

- **Adaptive Risk-Based Governance:** AI-IAs serve as adaptive, risk-based frameworks. Immersive realities require fluid regulatory models that evolve with continuous data capture, as AI systems are dynamic, not static.
- **Organizational Integration:** AI-IAs must be integrated into organizational processes. Public bodies, in particular, must be compelled to undertake self-auditing of AI systems.
- **Iterative Lifecycle Governance:** AI-IAs must be iterative, returned to at every new stage in the AI lifecycle. This supports the need for dynamic consent rather than one-off events.
- **Transparency and Accountability:** AI-IAs enable risk estimation, audit, and mitigation. Transparency, making processes accessible for inspection, is necessary for algorithmic responsibility.
- **Preventing "Ethics Washing":** The debate on AI ethics is at risk of being hijacked by corporate interests. To prevent "ethics washing," independent regulation and multi-stakeholder governance boards are required.

5.2.4 Specific Governance Recommendations

Based on the analysis, this study proposes specific governance actions:

For Biometric Governance:

- **Biometric Data Minimization:** Collect only necessary markers; prefer on-device processing and ephemeral storage.
- **Transparency by Design:** Provide real-time visual cues in AR/robotic interfaces during biometric capture.
- **Algorithmic Auditability:** Employ third-party auditing to check for bias and inferential integrity.
- **Global Ethics Harmonisation:** Design a Biometric Governance Charter, perhaps with UN/IEEE, to harmonize ethical benchmarks.

For Robotics Governance:

- **Ethical Robotics Charter:** Set global standards for moral design principles and social accountability.
- **Mandatory AI–Robotics Auditing:** Institutionalize lifecycle auditing for intent alignment, bias detection, and human-machine feedback loops.
- **Explanable Robotics Protocols:** Compel developers to generate interpretable algorithmic logs.
- **Chain-of-Accountability Registers:** Mandate registers mapping each decision layer to address distributed responsibility.
- **Mandatory Black-Box Recording:** Require recording of algorithmic reasoning and environmental data.
- **Social Impact Compensation Models:** Implement taxation or contribution schemes from automation deployments to finance reskilling programs.
- **Cross-Disciplinary Oversight Councils:** Include ethicists, engineers, and civil representatives in review processes.

This integrative vision ensures consent is an ongoing process, governance is proportional to data sensitivity, and regulation develops alongside technology, not in its wake.

6. Conclusion

A. Summary of the Proposed Framework

The study proposes a dual-pronged framework that integrates technical safeguards with policy-driven governance to set up a coherent model for ethical data management in the age of persistent augmented reality, robotics, and biometrics. On the technical side, the framework advocates for dynamic consent architectures, algorithmic transparency, AI impact assessments, and tiered data governance models that ensure proportional accountability and user agency. Complementing these mechanisms, the policy dimension calls for multi-stakeholder oversight, international harmonization of biometric laws, and human-rights-based regulation that evolves alongside technological affordances. The synergy between these dimensions underlines the idea that no ethical system will be effective if the technical and institutional architectures do not support each other-both translating moral intention into operational practice.

B. Long-Term Societal Implications

The long-term convergence of AR, robotics, and biometric analytics will reshape the boundary between public and private spaces by rendering everyday environments sites of constant perceptual data exchange. As spatial computing becomes integral to urban life, questions of visibility, consent, and surveillance will shape how societies negotiate privacy and collective security. This cultural evolution will require new social norms for data etiquette, much as industrial society once developed rules for traffic or the workplace. Public trust will become the crucial currency of this cohabitation in digital-physical space: without confidence in good governance, even groundbreaking technologies will face societal resistance and moral repudiation. The legitimacy of immersive systems thus depends less on their efficiency than on their harmony with democratic values of dignity, autonomy, and justice.

C. A Call to Action

Governance of emerging immersive technologies cannot afford to be reactive; proactive collaboration by developers, policymakers, ethicists, and civil society is called for. Technologists must embed moral reasoning within design; legislators must craft adaptive and globally consistent frameworks; and citizens must stay informed participants in shaping the ethical trajectory of innovation. Choices today—about how data are captured, interpreted, and governed—will determine not only the safety of digital ecosystems but the moral architecture of the societies that will live in them. Let this serve as a common call to make sure technological progress moves in a way that furthers human rights, transparency, and trust, while steering the future of augmented and robotic intelligence toward one that is truly responsible and respectful of rights.

References

1. Raja, U. S., & Al-Baghli, R. (2022). *Ethical Issues in Virtual Reality Today and Innovating Responsibly: Ethical Questions for AI in Early Childhood.* Journal of Media Ethics, 37(4), 215–229. https://doi.org/10.1080/08900523.2022.2056710
2. Jawalkar, S. K. (2021). *Ethical Horizons in Immersive Technologies: Challenges in AR/VR Governance.* International Journal of Ethics in Digital Society, 5(2), 112–130.
3. Braun, V., & Clarke, V. (2006). *Using Thematic Analysis in Psychology.* Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
4. European Union. (2016). *General Data Protection Regulation (GDPR).* Official Journal of the European Union, L119.
5. European Commission. (2024). *Artificial Intelligence Act.* COM(2021) 206 final — Regulation of the European Parliament and of the Council on Artificial Intelligence.
   U.S. Congress. (2018). *California Consumer Privacy Act (CCPA).* California Civil Code §§1798.100–1798.199.
6. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2022). *Ethically Aligned Design, First Edition.* IEEE Standards Association.
7. Floridi, L., & Cowls, J. (2019). *A Unified Framework of Five Principles for AI in Society.* Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1
8. Bryson, J. J., Diamantis, M. E., & Grant, T. D. (2017). *Of, for, and by the People: The Legal Lacuna of Synthetic Persons.* Artificial Intelligence and Law, 25(3), 273–291.
9. Coeckelbergh, M. (2020). *AI Ethics.* MIT Press.
10. European Parliamentary Research Service (EPRS). (2023). *AI and Robotics: Ethical and Legal Frameworks for the EU.* Brussels: European Parliament.
11. Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and Societal Implications of Algorithms, Data, and AI: A Roadmap for Research.* Nuffield Foundation.
12. Crawford, K., & Paglen, T. (2021). *Excavating AI: The Politics of Images in Machine Learning Training Sets.* International Journal of Cultural Studies, 24(6), 931–951.
13. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). *The Ethics of Algorithms: Mapping the Debate.* Big Data & Society, 3(2). https://doi.org/10.1177/2053951716679679
14. Borenstein, J., Herkert, J. R., & Miller, K. W. (2017). *The Ethics of Autonomous Cars.* The Atlantic Council Journal of Ethics in Engineering, 1(2), 1–15.
15. Calo, R. (2015). *Robotics and the Lessons of Cyberlaw.* California Law Review, 103(3), 513–563.
16. Lin, P., Abney, K., & Bekey, G. (2020). *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence.* Oxford University Press.
17. Zuboff, S. (2019). *The Age of Surveillance Capitalism.* PublicAffairs.
18. Raji, I. D., & Buolamwini, J. (2019). *Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products.* Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 429–435.
19. United Nations Educational, Scientific and Cultural Organization (UNESCO). (2021). *Recommendation on the Ethics of Artificial Intelligence.* Paris: UNESCO.