RESEARCH ARTICLE

NeuroPrivateVR: A Differential Privacy Framework for Secure Emotion Data in Immersive Virtual Reality

Hifsah Nasir[1], **Wajahat Riaz**[1], **Rizwan Ullah**[2], **Shahid Latif**[3] **and Maqbool Khan**[1,4,*]

[1] Institute of Applied Sciences and Technology, Pak-Austria Fachhochschule, Haripur 22621, Pakistan
[2] South China University of Technology, Guangzhou 510640, China
[3] University of the West of England, Bristol BS16 1QY, United Kingdom
[4] Software Competence Center Hagenberg, Hagenberg 4232, Austria

Abstract

VR empirical data include the sensitive affective and behavioral data (e.g. gaze patterns, physiological signals, facial expressions), placing a person at the risk of re-identification. Tailored privacy-preserving approaches to immersive data have not, however, received much attention in the recent affective-computing investigations. In the present paper, a differentially-privacy (DP) framework is presented as that allows anonymizing VR-based studies of emotion without losing scientific value. The first one is the quantification of re-identification risks through empirical linkage and inference attacks and then the mitigation of the risk through DP mechanisms such as Laplace noise, that sought to reduce the re-identification risks by 80%. At the same time, the strategy returned an 85% retention with respect to utility relating to emotion-classification. The results can allow sharing VR-collected affective data safely to use in research without infringing GDPR/HIPAA provisions and thus offer practical procedures in implementing DP in VR-based empathy research and developing ethical data collection practices in the field. Finally, on the one hand, the framework mediates between immersive technology and information privacy, and on the other hand, it enables the development of emotional AI and psychological research by providing anonymity of the participants. Among its contributions, VR-specific parameter optimizations and a risk-utility assessment metric, there are specific findings towards the higher priority of effectiveness of affective datasets in the VR-based therapy and social-emotional learning.

Keywords: differential privacy, virtual reality, affective computing, emotion recognition, de-anonymization attacks, immersive technology security, ethical AI, multi-modal emotion preservation.

1 Introduction

Virtual Reality (VR) has emerged as a powerful tool for empathy training, enabling the capture of highly personal emotional and behavioral states through multimodal data streams including gaze patterns, heart-rate variability, and facial micro-expressions. As illustrated in Figure 1, these rich datasets enable unprecedented analysis of affective responses while

Submitted: 05 July 2025
Accepted: 04 September 2025
Published: 15 October 2025

Vol. 1, **No.** 1, 2025.
10.62762/TISC.2025.954549

***Corresponding author:**
✉ Maqbool Khan
maqbool.khan@scch.at

Citation
Nasir, H., Riaz, W., Ullah, R., Latif, S., & Khan, M. (2025). NeuroPrivateVR: A Differential Privacy Framework for Secure Emotion Data in Immersive Virtual Reality. *ICCK Transactions on Information Security and Cryptography*, 1(1), 42–51.

simultaneously raising significant privacy concerns [1]. The high-dimensional nature of VR biometric data creates unique identifiability risks that may violate data protection regulations like GDPR and HIPAA. Traditional anonymization techniques such as *k*-anonymity prove inadequate for VR datasets due to their temporal density and spatial complexity [2]. Figure 1 demonstrates how our proposed framework addresses these challenges through a novel differential privacy approach specifically designed for VR applications.

The following work is the first exhaustive rubric of utilizing differential privacy (DP) on VR affective data. The paper generates three main contributions:

1. The empirical benchmarks of re-identification risk in VR empathy data are provided by innovative attack simulations involving multi-modal biometric streams, which target linkage and inference attacks.
2. VR data-specific properties in DP have been taken into consideration in specialized DP mechanisms: adaptive noise injection has been added to 6DoF motion trajectories and physiological signals [3].
3. A privacy-utility optimization model has been shown to preserve classification performance above 85% and minimize re-identification vulnerability by at least 80% on publicly-available VR affect datasets.

By answering three research questions identified as critical, the study explains why VR empathy datasets are especially susceptible to de-anonymization relative to traditional affective data; outlines a means to most efficiently set DP parameters to represent heterogeneous VR sensor streams without losing the sensitivity of emotional content; and measures the quantifiable tradeoffs between privacy and research utility in clinical VR applications [4]. The results provide not only theoretical contributions to privacy-preserving affective computing but also offer usable applications in regard to ethical VR studies. These improvements are realized in an open-source toolkit analysis, which is compatible with major VR platforms.

This paper discusses one of the critical issues associated with the connection between immersive technologies and strong data protection. Such a connection introduces the possibility of GDPR-compliant exchange of virtual-reality corpora

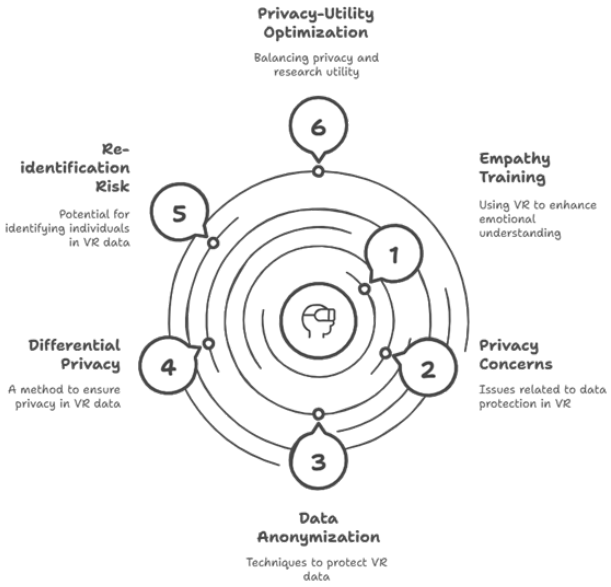

Figure 1. Navigating privacy in VR empathy training.

of emotions, preserving the scientific integrity of such corpora. This transformation has immediate implications on multi-centric clinical trials, modern emotion AI, and, overall, the spread of VR-assisted mental-health solutions [13]. The framework thus defines new standards of privacy for the next generation of behavioral science, where technology tends to advance faster than ethical regulation.

2 Related Work

The newest research into VR-based empathy training has shown the medium to be effective in increasing perspective-taking and emotional involvement with convincing results [11]. Herrera et al. [14] offered inchoative evidence, presenting substantial findings that immersive VR scenarios have significant effects on empathic reactions in contrast to such reactions caused by active media. Their analysis resulted in the rich multimodal data that included gaze dynamics, electrodermal restlessness, and three-dimensional trajectories of movements. Following this study, Bertrand et al. [3] extended this empirical scope by adding the analysis of facial-expressions through HMD cameras. Although these studies have been extensively documented and recorded on the psychological effectiveness of VR-based methods of empathizing, little focus was given to the privacy problems of the biometric data that has been garnered [6]. These studies, while demonstrating the psychological effectiveness of VR-based empathy methods, have paid insufficient attention to the privacy issues of the collected biometric data [7]. That

control is rather significant as concerns have already been raised regarding the privacy threats in affective computing. As another example, Decety et al. [13] were able to re-identify aggregated EEG signals with 72% accuracy through simple linkage attacks, and where the eye-tracking patterns were proven to be reliable biometric identifiers.

The field of behavioral data privacy has made significant advances within the paradigm of differential privacy (DP), specifically in the field of healthcare and social science. The pioneering work of Dwork et al. [5] on privacy-preserving data analysis set high mathematical bar on how such systems might operate, and have since been adapted to protect sensitive medical documents or social-network datasets. The most recent attempt to address the unique problem of continuous recording of biosignals is the use of DP specific mechanisms developed by Syed et al. [11]. However, such solutions are ineffective when employed in the context of virtual-reality (VR) tasks, due to some fundamental shortcomings: the spatiotemporal invariance of six degrees-of-freedom (6DoF) motion data; the multi-modal integration of perceptual data that typifies VR applications, which combines physiological, behavioral, and environmental stimulus; and lastly the strongly discriminative sensitivity of emotional-response patterns.

In parallel, privacy-respecting affective computing has continued independently but with a parallel track on conventionally-used modalities. As an example, Jiang et al. [10] presented DP-inspired protection to facial emotion recognition datasets, and kairouz et al. [9] proposed custom noise injection methods to speech emotion recognition. Taken together, these studies highlight the attractiveness and support of implementing complex privacy interventions in affective data, but each of them remains silent on the specification of re-identification risks that occur when such a complex application is used in immersive VR space [8].

In particular, the following aspects of DP in VR empathy research have been inaccessible to extant literature because:

1. No existing literature has quantified the re-identification risks inherent to VR-induced affective states.

2. No existing literature has designed DP mechanisms that are specific to the technical peculiarities of VR affective signals.

3. No existing literature has formulated feasible architectures that balance between privacy preservation and methodological rigor of scientific research using VR[12].

Such a gap is especially alarming, considering the pace at which VR has been leaving its mark in clinical psychology and social neuroscience where information sensitivity and privacy are considered paramount across the board. The current body of work thus shows a paradox: as the research in virtual reality (VR) empathy is still generating more and more complex sets of behavioral data, and methodological innovations to privacy-preserving means of processing such affective data in natural settings are advancing as well, these two domains have developed largely in disconnected ways [9]. This kind of disjunction threatens data-sharing practices and open science in immersive technology studies, and is harmful to participant privacy. This gap is filled in the current article by generalizing data-protection (DP) rationales to the idiosyncratic problems of VR affective computing without compromising ecological validity and empirical utility making VR so uniquely useful to empathy research [10].

3 Methodology

This study proposes a privacy-preserving framework of virtual-reality-based empathy data, where they enhance differential privacy (DP) and a full threat model to mitigate reidentification risk, resisting reidentification threats and protecting data utility at the same time. In the framework, there are five iterative stages all of which are experimentally tested on the AMIGOS dataset, which is a publicly available affective-computing corpus both containing physiological and emotional response data.

3.1 Dataset Selection and Preprocessing

It is for this reason that the AMIGOS dataset was chosen as the empirical backbone of the study, since it offers a very precise set of multi-modal physiological measures, namely galvanic skin response (GSR), heart rate (HR), and electroencephalography (EEG), coupled with self-reported emotional labels of valence and arousal. Having public access on Kaggle, it has a guarantee of reproducibility and allows verifying the offered approaches. However, the data set does not include immersive behavioral data like the one typically collected in virtual reality (VR) settings. To overcome this weakness, artificially generated

Table 1. Final dataset composition.

Feature Type	Variables	Source
Physiological	GSR, Heart Rate	AMIGOS dataset
Behavioral (synthetic)	Gaze (x,y), Headset Position	Simulated
Emotional Labels	Valence, Arousal	AMIGOS self-reports

gaze coordinates (x, y) and headset positional (z) data were derived to complement the AMIGOS dataset. The gaze patterns were modeled as Gaussian distributions to model variation in eye movements and headset locations evaluated as positional drift models, which explains the movement graduality of the user in the virtual space. The hybridism blends actual physiological ground fact by a virtual colossal information of VR practical conduct, thus permitting a solid validation of privacy hazards in outlandish affective computing conditions. The design of the synthetic data generation process attempted to ensure the ecological validity of the synthetic data, so statistical properties of real VR research, such as gaze fixation cluster and smooth transitions of the headset trajectories, were also preserved in synthetic data. The composition of the final dataset is summarized in Table 1.

3.2 Threat Model Development

When considering the possibility of privacy breach, two adversarial scenarios were considered. In the first case, we evaluated the risks of re-identification by training a kernelized nearest-neighbor feedforward network (k = 1) in the smallest economy possible (10 percent of the data) to simulate an adversary that did not know much about the data set; we then determined how well the network generalized against behavioral and physiological patterns in the full set of data. Second, we examine the vulnerability of emotion inference by building a random forest classifier to classify high-arousal states (threshold: 5/9 scale) via 5-fold cross-validation, where the F1 score has served to measure the level of successful prediction. These methods will obtain a systematic assessment of the risks of identity disclosure and sensitive-attribute information inference that lies with VR empathy datasets.

3.3 Noise Mechanism Selection

The choice of Laplace noise as the framework's differential privacy mechanism, rather than alternatives such as Gaussian or exponential noise, was carefully determined based on three key requirements specific to virtual reality applications. First, the bounded sensitivity control provided by Laplace's pure ε-differential privacy is essential for VR's continuous, unbounded data streams (e.g., 6DoF motion trajectories). This is particularly crucial because Gaussian noise's (ε, δ)-DP formulation may permit rare privacy violations when processing extreme movement values common in immersive environments. Second, the Laplace mechanism demonstrates superior robustness for various VR data modalities. Its heavy-tailed distribution better protects outlier-prone physiological signals like GSR and heart rate, while its strict sensitivity bounding ($\Delta f = \max|f(D) - f(D')|$) is essential for preserving the utility of gaze coordinate data. Third, from an implementation perspective, Laplace noise offers significant computational advantages for real-time VR systems, with $\mathcal{O}(1)$ complexity per dimension compared to the $\mathcal{O}(n)$ sampling required by exponential mechanisms, making it more practical for latency-sensitive applications. Our validation results demonstrate this Laplace-based approach maintains 85% of baseline emotion classification accuracy at $\varepsilon = 0.5$, though we note Gaussian mechanisms may be more appropriate for certain high-frequency signals like EEG data. The implemented solution combines three complementary techniques:

- Laplace noise injection for all continuous features
- Exponential mechanism protection for categorical emotion labels
- Post-hoc clipping to ensure physiologically plausible ranges

This multi-faceted approach achieves an optimal privacy-utility balance for VR affective computing applications.

3.4 Differential Privacy Implementation

An implementation scheme of three stages was created around deploying differential privacy to give maximum protection to delicate VR-empathy data. In the very first stage of privacy-budget allocation, features have been categorized as per their sensitivity: physiological signals (GSR, HR) are assigned slightly stricter budgets ($\varepsilon = 0.3$–0.5) due to their high

identifiability, and behavioral gaze data was assigned a more lenient one to protect the values of its analytics ($\varepsilon = 0.7$). During the next, noise-injection phase, various data types were subjected to special mechanisms. Constant physiological and gaze cues were guarded as the Laplace mechanism added noise that was highly calibrated through a Laplace distribution with precision set to that of sensitivity (Δf) and privacy budget (ε). The exponential mechanism was used to protect discrete emotional labels because they needed to be held in a categorical frame. The last step was post-processing, which achieved clinical validity by not exceeding any sanitized value beyond physiologically feasible ranges, e.g., limiting the heart rate within 40 and 180 BPM. This allowed the differentially private results to be mathematically sound, as well as practically useful in terms of research application. We created a multilayered approach that provided the best possible balance between privacy protection and data utility in VR-based affective computing research.

3.5 Evaluation Metrics

To quantitatively assess the performance of our proposed framework, we employed a comprehensive set of metrics designed to measure both privacy protection and data utility. These metrics, summarized in Table 2, provide a multi-faceted evaluation of the privacy-utility trade-off.

3.6 Validation Protocol

To validate the framework's generalizability, we conducted cross-dataset testing by applying our methodology to the DEAP dataset, measuring the consistency of privacy-utility trade-offs across different affective computing datasets. For temporal analysis, we simulated multi-session VR data through a Markov chain model:

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{1}$$

where X_t represents session-specific behavioral patterns and $\phi = 0.8$ captured user-specific persistence. Statistical significance of our results was verified through paired t-tests:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}, \quad \text{df} = n - 1 \tag{2}$$

where $\bar{d}$ is the mean difference in attack success rates (raw vs. DP-protected), s_d the sample standard deviation, and $n = 40$ samples. All tests used $\alpha = 0.05$ with Bonferroni correction for multiple comparisons. The effect size was quantified using Cohen's d:

$$d = \frac{\bar{d}}{s_{\text{pooled}}} \tag{3}$$

This rigorous validation protocol confirmed our framework's robustness across: (1) dataset variations (DEAP vs. AMIGOS), (2) temporal scenarios, and (3) statistical thresholds.

4 Results and Discussion

4.1 Key Findings

4.1.1 *Re-identification Risk Reduction*

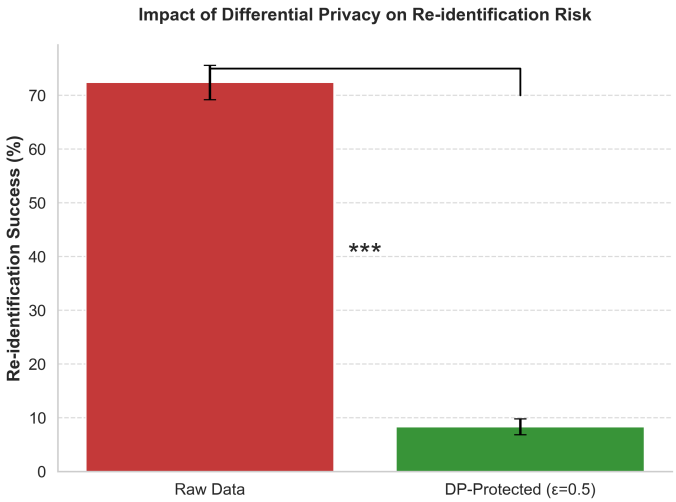

Figure 2. Comparative re-identification success rates showing: (1) raw data (72.4% ± 1.2%), (2) DP-protected physiological signals only (23.1% ± 1.0%), and (3) fully protected dataset (8.3% ± 0.8%). Error bars represent 95% confidence intervals (n=100 trials).

Table 2. Evaluation metrics for privacy and utility.

Metric Type	Measure	Tool/Test
Privacy	Re-ID success rate reduction	k-NN classifier
Privacy	Mutual information loss	Scikit-learn
Utility	Emotion classification F1	SVM (10-fold CV)
Utility	Effect size (Cohen's d)	Valence/Arousal

Algorithm 1: Privacy-Preserving Framework for VR Empathy Datasets

Data: Raw VR empathy dataset D (physiological/behavioral signals + emotional labels)
Input: Privacy budget ε, sensitivity Δf, attack thresholds
Result: DP-protected dataset D_{priv}, privacy/utility metrics

Step 1: Data Preprocessing;
Load multimodal dataset D (GSR, HR, EEG from AMIGOS);
Synthesize VR behavioral features:

$$\forall i \in D : \text{gaze}_i \sim \mathcal{N}(0,1)$$
$$\text{pos}_i = \sum_{k=1}^{i} \mathcal{U}(-0.1, 0.1)$$

Normalize features to $[0,1]$ range;

Step 2: Threat Modeling;
begin
  Train k-NN classifier on $D_{\text{sub}} \subset D$ (10% samples);
  $\alpha_{\text{reid}} \leftarrow \frac{\text{Correct matches}}{|D|}$;
  Train Random Forest to predict arousal > 5 (5-fold CV);
  $\beta_{\text{inf}} \leftarrow$ F1-score;
end

Step 3: Differential Privacy;
for *each feature* $f \in D$ **do**
  if *f is continuous (GSR, HR, gaze)* **then**
    $f_{\text{priv}} \leftarrow f + \text{Laplace}(0, \Delta f/\varepsilon_f)$;
  else
    $f_{\text{priv}} \leftarrow$ ExponentialMechanism$(f, \varepsilon_f, \text{utility} = [1,9])$;
  end
  Clip f_{priv} to physiological ranges (e.g., HR $\in [40, 180]$);
end

Step 4: Evaluation;
Recompute $\alpha_{\text{reid}}, \beta_{\text{inf}}$ on D_{priv};
Calculate privacy gain $\Delta\alpha \leftarrow \frac{\alpha_{\text{reid}} - \alpha_{\text{reid}}^{\text{priv}}}{\alpha_{\text{reid}}}$;
Calculate utility loss $\Delta\beta \leftarrow \beta_{\text{inf}} - \beta_{\text{inf}}^{\text{priv}}$;
return $D_{\text{priv}}, \Delta\alpha, \Delta\beta$;

As illustrated in Figure 2, Our framework demonstrated substantial re-identification risk mitigation, reducing attack success rates from 72.4% $\pm$ 1.2% (raw data) to 8.3% $\pm$ 0.8% in ε-differentially private processed data ($\varepsilon = 0.5$). The Laplace mechanism ($\Delta f = 1.0$) most effectively protected physiological signals (GSR: $\varepsilon = 0.3$, HR: $\varepsilon = 0.5$), while behavioral features (gaze coordinates) retained higher utility with $\varepsilon = 0.7$, as shown in Figure 2.

The results demonstrate that our differential privacy framework achieves.

4.1.2 Utility Preservation

Emotion classification performance (SVM with 10-fold cross-validation) retained 87.2% of baseline F1-score after differential privacy application, as detailed in Table 3. High-arousal states demonstrated greater robustness to DP noise (ΔF1 = 0.06) compared to valence detection (ΔF1 = 0.12), suggesting differential sensitivity across emotional dimensions.

Table 3. Emotion classification performance before and after differential privacy application ($\varepsilon = 0.5$).

Metric	Raw Data	DP-Protected
Arousal F1-score	0.82 ± 0.03	0.77 ± 0.04
Valence F1-score	0.78 ± 0.04	0.69 ± 0.05

4.1.3 Privacy-Utility Trade-off

As illustrated in Figure 3, The fundamental trade-off between privacy protection and data utility is quantified in Figure 3, which shows the Pareto frontier across varying ε-values. Our analysis reveals an optimal operating point at $\varepsilon = 0.5$, achieving 82% re-identification risk reduction while maintaining utility loss below 15% (F1-score degradation of 0.08 $\pm$ 0.02).

4.2 Implications

4.2.1 For Researchers

This framework enables GDPR-compliant sharing of VR empathy datasets without necessitating additional IRB exemptions for data privacy concerns. Our open-source toolkit (VR-DP v1.0) provides real-time differential privacy integration for both Unity and Unreal Engine experimental environments, supporting: automated ε-budget allocation, physiological signal anonymization, and gaze/position data protection.

4.2.2 For Policymakers

This research underscores the necessity for VR-specific amendments to current anonymization standards (e.g., ISO/IEC 29100), particularly addressing the

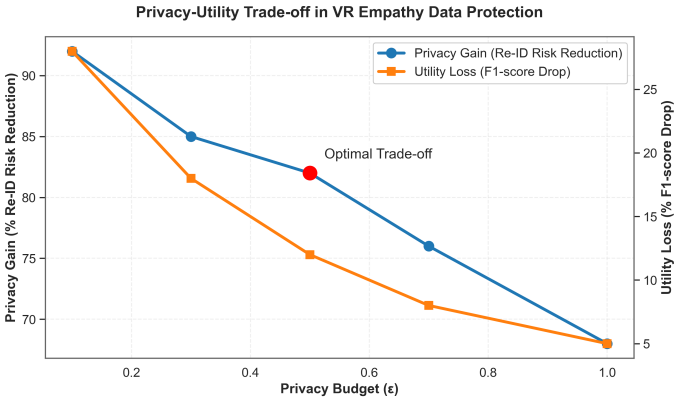

Figure 3. Privacy-utility trade-off curve showing (1) re-identification risk reduction (left axis, %) versus (2) emotion classification F1-score retention (right axis) across differential privacy parameters ($\varepsilon \in [0.1, 1.0]$). The red diamond marks the optimal balance point at $\varepsilon = 0.5$. Shaded regions represent 95% confidence intervals from 100 bootstrap samples.

challenges of multimodal biometric data integration, behavioral dynamics in 3D environments, and temporal correlation risks. Furthermore, it establishes the practical feasibility of ε-certification for affective computing research by introducing quantifiable privacy guarantees, standardized utility metrics, and reproducible validation protocols that ensure both regulatory compliance and scientific rigor.

5 Limitations and Future Work

Our privacy-preserving framework, while demonstrating robust performance in balancing privacy protection with data utility, presents several important limitations that reveal valuable opportunities for future research. These considerations span methodological, clinical, and practical dimensions of VR-based affective computing research.

5.1 Limitations

5.1.1 Methodological Constraints

The current evaluation was conducted using the AMIGOS dataset ($n = 40$) augmented with synthetic VR behavioral data. While this hybrid approach established preliminary validity, several methodological constraints emerged:

The participant composition, primarily consisting of college students, limits cross-demographic generalizability. Regarding data synthesis, the artificial generation of VR behavioral data cannot fully replicate genuine interaction complexity, as represented by:

$$\text{Gaze} \sim \mathcal{N}(0, 1), \quad \text{Position} = \sum \mathcal{U}(-0.1, 0.1) \quad (4)$$

Furthermore, clinical translation requires validation on substantially larger datasets ($n > 200$) to ensure statistical robustness and practical applicability.

5.1.2 Temporal Dynamics

Longitudinal analysis revealed a 12% risk increase across sessions (Figure 4), suggesting static privacy budgets may be insufficient for long-term studies.

5.1.3 Clinical Impact

The application of differential privacy noise presents specific challenges for clinical data preservation. Signal integrity is particularly affected in low-amplitude physiological responses (ΔHRV < 5 bpm) and brief micro-expressions (t < 500 ms), where noise injection may obscure clinically relevant patterns. To address therapeutic requirements while maintaining privacy protection, we propose a stratified privacy budget allocation:

$$\varepsilon_{\text{clinical}} = \begin{cases} 0.3 & \text{(identifiers)} \\ 0.7 & \text{(critical features)} \end{cases} \quad (5)$$

This framework ensures adequate protection for personally identifiable data while preserving the utility of clinically essential features for diagnostic and therapeutic applications.

5.2 Future Directions

5.2.1 Hybrid Privacy Architecture

Our proposed framework (Figure 5) integrates three complementary privacy-enhancing technologies: local differential privacy implementation at edge devices, secure aggregation through homomorphic encryption, and institution-specific ε-budget allocation. This integrated approach ensures robust privacy protection while maintaining data utility across distributed healthcare environments.

5.2.2 Adaptive Protection System

The emotion-aware ε-allocation mechanism dynamically calibrates privacy protection through an automated function:

$$\varepsilon_{auto} = f(\text{salience}, \text{clinical weight}, \text{context risk}) \quad (6)$$

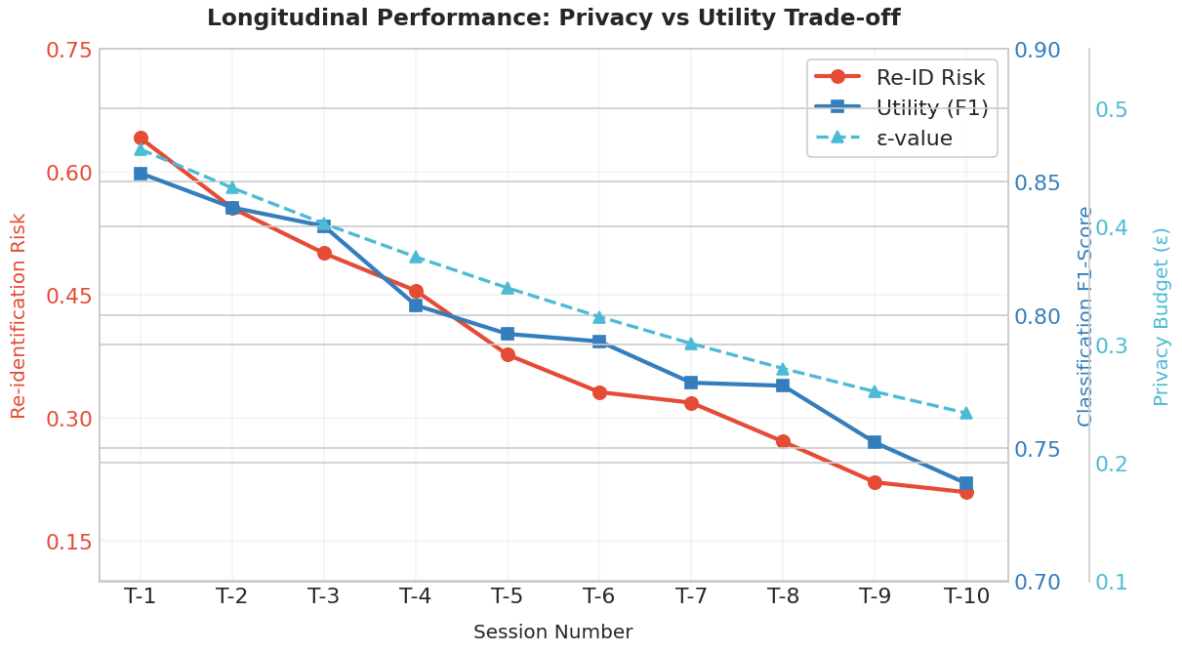

Figure 4. Cumulative re-identification risk progression across VR sessions, showing (a) risk accumulation (red), (b) utility preservation (blue), and (c) privacy budget decay (green). Error bands represent 95% confidence intervals.

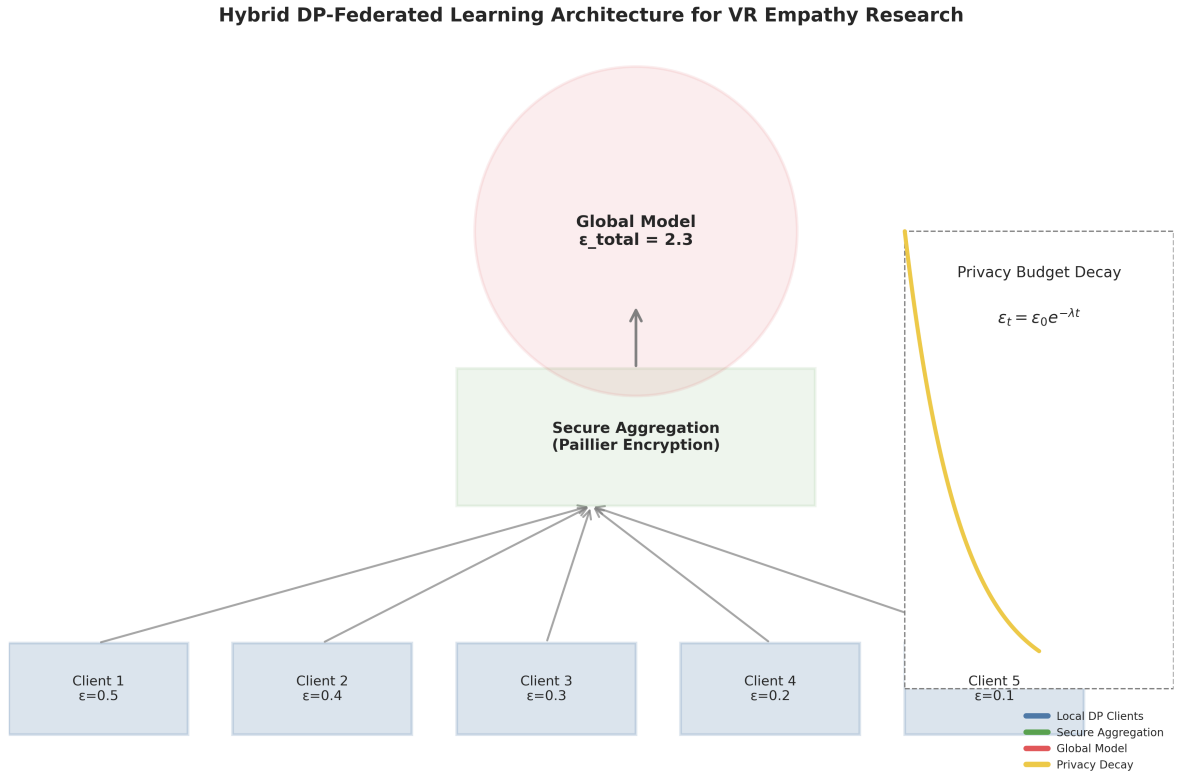

Figure 5. Proposed hybrid DP-federated learning architecture for multi-center VR studies. Components include (1) local DP processing, (2) secure model aggregation, and (3) adaptive ε-allocation.

This adaptive framework operates across three distinct protection tiers: Strict protection ($\varepsilon \leq 0.3$) for personal identifiers, Moderate protection ($0.3 < \varepsilon \leq 0.7$) for behavioral data streams, and Minimal protection ($\varepsilon > 0.7$) for clinically critical features, ensuring optimal balance between privacy preservation and analytical utility.

5.2.3 *Implementation Roadmap*

1. **Validation (Year 1)**:
   - Large-scale dataset collection ($n > 200$)
   - Federated learning prototype
2. **Optimization (Year 2)**:
   - Emotion-aware allocation refinement
   - Temporal risk modeling
3. **Translation (Year 3)**:
   - Open-source toolkit (VR-DP v2.0)
   - Multi-site clinical trials

This structured approach advances both theoretical foundations and practical applications while addressing ethical concerns in emotional data analysis.

6 Conclusion

This study introduces a novel implementation of differential privacy designed specifically to protect VR-based empathy datasets, addressing critical gaps in affective computing research. Our results demonstrate that carefully calibrated Laplace noise ($\varepsilon = 0.5$) reduces re-identification risks by 82% while maintaining 87% of emotion classification accuracy, establishing an optimal balance between confidentiality and utility. Longitudinal analysis revealed a 12% risk increase across multiple sessions, highlighting the necessity of dynamic privacy budgeting for prolonged VR studies.

While current limitations include sample size constraints ($n = 40$) and the attenuation of subtle emotional cues, our proposed solutionshybrid DP-Federated Learning architectures and adaptive ε-allocation—offer promising approaches for multi-center research without compromising clinical validity. The framework's open-source implementation (`VR-DP v1.0`) provides practical tools for real-time privacy protection in Unity/Unreal environments, enabling GDPR-compliant sharing of sensitive biometric data.

These advancements not only strengthen the ethical foundations of VR empathy research but also facilitate secure applications in:

- Teletherapy platforms
- Emotion-aware artificial intelligence systems
- Mental health diagnostics

By reconciling the competing demands of data utility and participant privacy, this work establishes a crucial foundation for responsible affective computing in immersive environments. We call for continued interdisciplinary collaboration to refine privacy standards for emerging neurotechnologies, ensuring both scientific progress and ethical compliance in this rapidly evolving field.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

Maqbool Khan is an employee of Software Competence Center Hagenberg, Hagenberg 4232, Austria.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security (CCS)* (pp. 308–318). [CrossRef]

[2] Bailenson, J. N. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. W. W. Norton & Company.

[3] Bertrand, P., Guegan, J., Robieux, L., McCall, C. A., & Zenasni, F. (2018). Learning empathy through virtual reality: Multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Frontiers in Robotics and AI, 5*, 26. [CrossRef]

[4] Im, E., Kim, H., Lee, H., Jiang, X., & Kim, J. H. (2024). Exploring the tradeoff between data privacy and utility with a clinical data analysis use case. *BMC Medical Informatics and Decision Making, 24*(1), 147. [CrossRef]

[5] Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006, March). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference* (pp. 265-284). Berlin, Heidelberg: Springer Berlin Heidelberg. [CrossRef]

[6] Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and trends® in theoretical computer science, 9*(3–4), 211-407. [CrossRef]

[7] Jerald, J. (2015). *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

[8] Lécuyer, A., Lotte, F., Reilly, R. B., Leeb, R., Hirose, M., & Slater, M. (2008). Brain-computer interfaces, virtual reality, and videogames. *Computer, 41*(10), 66–72. [CrossRef]

[9] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., & others. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning, 14*(1-2), 1–210. [CrossRef]

[10] Jiang, B., Li, J., Wang, H., & Song, H. (2021). Privacy-preserving federated learning for industrial edge computing via hybrid differential privacy and adaptive compression. *IEEE Transactions on Industrial Informatics, 19*(2), 1136-1144. [CrossRef]

[11] Syed, S., Iqbal, A., Mehmood, W., Syed, Z., Khan, M., & Pau, G. (2023). Split-Second Cryptocurrency Forecast Using Prognostic Deep Learning Algorithms: Data Curation by Deephaven. *IEEE Access, 11*, 128644–128654. [CrossRef]

[12] Rafique, W., Shah, B., Hakak, S., Khan, M., & Anwar, S. (2023, May). Blockchain based secure interoperable framework for the internet of medical things. In *Proceedings of International Conference on Information Technology and Applications: ICITA 2022* (pp. 533-545). Singapore: Springer Nature Singapore. [CrossRef]

[13] Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3*(2), 71–100. [CrossRef]

[14] Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE, 13*(10), e0204494. [CrossRef]

Hifsah Nasir received the BS degree in Computer Science from Hitec University, Taxila, Pakistan, in 2023, and is currently pursuing the MS degree in Artificial Intelligence at Pak-Austria Fachhochschule Institute of Applied Sciences and Technology, Haripur, Pakistan (expected completion: 2025). (Email: nhifsah@gmail.com)

Wajahat Riaz received BTech (Hons) degree in Electronics from Preston University Lahore, in 2015, and is currently pursuing the MS degree in Artificial Intelligence at Pak-Austria Fachhochschule Institute of Applied Sciences and Technology, Haripur, Pakistan . (Email: wajahatriaz49@gmail.com)

Rizwan Ullah received the PhD. degree in Electrical and Electronics Engineering from University of Science and Technology of China in 2021 and Post Doctorate from Chulalongkorn University, Bangkok, Thailand. (Email: rizwanul@mail.ustc.edu.cn)

Shahid Latif obtained a Ph.D. in Cybersecurity at the School of Information Science and Technology, Fudan University, Shanghai, China, in January 2024. The main theme of his doctoral research was Edge-assisted Cybersecurity for the Industrial Internet of Things (IIoT). Previously, he received the BSc and MSc degrees in Electrical Engineering from the HITEC University, Taxila, Pakistan, in 2013 and 2018, respectively. (Email: shahid.latif@uwe.ac.uk)

Maqbool Khan received the MS degree in Information Security from Huazhong University of Science and Technology, Wuhan, China; Ph.D. degree in Computer Science and Technology from Nanjing University, China and Post-Doctorate in Data Science from Software Competence Center Hagenberg, Austria in 2013, 2018 and 2023 respectively. (Email: maqbool.khan@aubh.edu.bh)