Who Is There With Me? User Identity, Avatars, and Authenticity in Social XR

Chuyang Zhang[1], Hyeongil Nam[1] and Kangsoo Kim[1,*]

[1]University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada

Abstract

In social extended reality (XR), avatars are often perceived as representations of real users, creating strong expectations of authenticity and identity consistency. However, existing XR authentication methods primarily address access control and do not explain how verified identity translates into perceived identity legitimacy, particularly in asynchronous, agent-mediated interactions. We propose a conceptual framework that reuses authentication-derived biometric and trait-based signals as shared user profile representations for both identity verification and avatar behavior generation. By linking verified identity with identity-consistent verbal and nonverbal expression, the framework maps authentication signals to perceived identity authenticity. We outline key research challenges and future directions to guide implementation and empirical validation.

Keywords

Extended Reality, Virtual Reality, Social XR, User Identity and Profile, Avatar, Virtual Agent, Authentication

1. Introduction

Social extended reality (XR) systems increasingly rely on human-like embodied avatars to support remote interaction, enabling users to communicate through multimodal cues such as gaze, posture, gesture, and facial expression. Prior work has shown that avatar-mediated communication can strengthen social presence and perceived interpersonal connection compared to conventional remote collaboration tools [1, 2, 3].

As Social XR expands into asynchronous contexts where users interact across time zones or through delayed participation, agentic avatars are emerging to bridge temporal gaps [4]. Prior work has shown that asynchronous VR playback can improve immersion compared to conventional media, but still produces weaker engagement and reduced social presence when users are constrained to passive viewing [5, 6]. Recent systems demonstrate that AI-driven embodied avatars can reconstruct or simulate prior participants, improving engagement and perceived communication quality in asynchronous settings [7]. This evolution suggests that Social XR may increasingly adopt avatar-agent hybrid systems, where user identity persists beyond live participation. However, this shift introduces a fundamental tension: while persistent embodied agents may increase engagement and continuity of interaction, they also amplify risks of identity misuse, impersonation, and behavioral manipulation.

Unlike conventional AI assistants, embodied avatars in Social XR are often interpreted as representations of real people, creating strong expectations of authenticity and identity consistency. Prior research on avatar identity and authentication in VR highlights the difficulty of ensuring that a virtual representation genuinely corresponds to its claimed user, as visual and behavioral features can be manipulated for impersonation or spoofing [8]. Existing authentication approaches largely focus on access control (e.g., login or device verification) and do not address whether an avatar's observable behavior remains consistent with the registered identity — a gap noted in systematic reviews of immersive authentication methods [9]. This issue becomes more pronounced in agentic and asynchronous contexts, where users may interact with recordings, generated proxies, or simulated representations without clear assurances

SAXR CHI'26: ACM CHI 2026 Workshop on Shaping Future Human Connection: Social Augmentation through XR Technologies, April 13–17, 2026, Barcelona, Spain

*Corresponding author.

chuyang.zhang1@ucalgary.ca (C. Zhang); hyeongil.nam@ucalgary.ca (H. Nam); kangsoo.kim@ucalgary.ca (K. Kim)

0009-0009-7088-9953 (C. Zhang); 0000-0002-6017-8869 (H. Nam); 0000-0002-0925-378X (K. Kim)

© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-07-17

CEUR-WS.org/Vol-4226/paper09.pdf

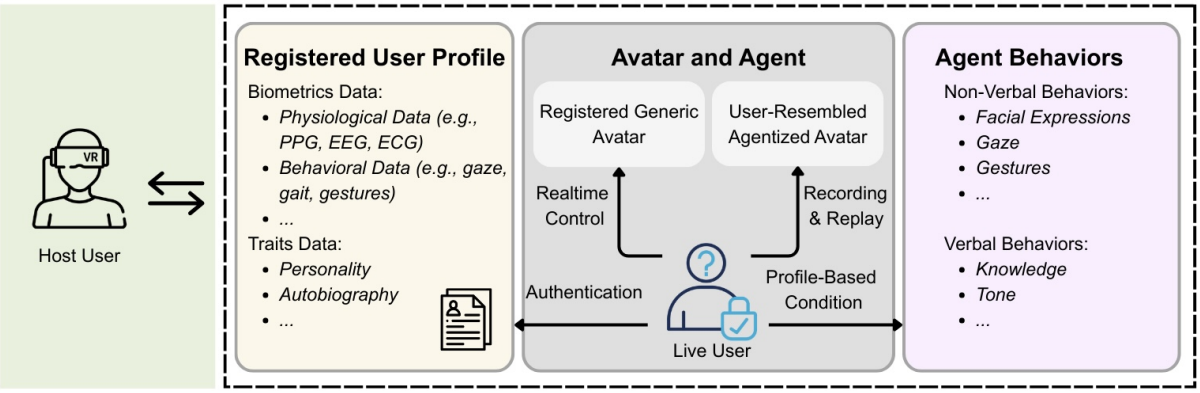

Figure 1: Overview of the proposed framework linking authentication-derived user profiles to avatar and agent behavior in synchronous and asynchronous Social XR.

of identity continuity.

At the same time, biometric and behavioral signals have demonstrated strong potential for user identification: physiological biometrics such as ECG, EEG, and PPG offer high identity specificity [10], while gaze, gesture, and movement patterns can be continuously captured through XR sensors [11, 12, 13]. These multimodal data streams are increasingly embedded in consumer XR systems. However, it remains unclear how system-level authentication confidence can be translated into users' perception of identity legitimacy — that is, whether the avatar they encounter genuinely corresponds to the individual it claims to represent [14, 15, 7]. In other words, while authentication mechanisms may verify identity internally, users lack perceptible cues that support the authenticity of the avatar in ongoing interaction.

In this paper, identity legitimacy refers to users' confidence that the avatar they encounter genuinely corresponds to its claimed individual, rather than to a manipulated, impersonated, or generatively altered representation. This construct differs from visual realism or general social presence; instead, it concerns whether authentication confidence is perceptually supported through identity-consistent behavioral expression across time.

We argue that Social XR lacks a systematic framework for linking authentication signals to perceived identity authenticity in agent-mediated interaction. To address this gap, we propose a conceptual model that reuses multimodal authentication-derived user profiles to both verify identity and condition avatar behavior. By mapping authentication-relevant signals to identity-consistent verbal and nonverbal expression, the framework raises new research questions about how user profiles can meaningfully and responsibly shape perceived identity consistency without unintended consequences. In this position paper, we discuss the prospects and challenges of designing such Social XR systems as avatars become persistent and agentic across temporal gaps.

2. Proposed Framework

2.1. Framework Design

Figure 1 illustrates an overview of our framework. Our framework supports avatar-mediated communication in both synchronous and asynchronous scenarios, where avatars function as persistent representations of registered users. In both settings, a host user interacts with avatars that are pre-registered and bound to specific user profiles derived from biometric, behavioral, and trait-based data. The selected modalities are limited to signals that can also be captured during interaction. This enables the system to authenticate the user based on incoming inputs in synchronous or asynchronous scenarios, while simultaneously personalizing avatar behaviors to reflect user-specific patterns.

In synchronous communication, the registered avatar is directly controlled by a live user. Authentication is required before control is granted. The system analyzes incoming interaction data, including the

HMD- and controller-tracked behavioral signals such as gaze and gesture, and compares them against the stored profile. When motion and interaction patterns alone are insufficient for reliable identification, physiological signals can serve as supplementary authentication factors.

In asynchronous or agent-mediated scenarios, the stored profile is reused not only for verification but also to condition avatar behavior. Authentication-derived representations inform both nonverbal patterns (e.g., gesture timing, movement variability) and verbal expression (e.g., tone, vocabulary). By integrating authentication and behavior generation through a shared profile, the framework supports identity continuity across temporal gaps.

2.2. User Profiling

In our framework, user profiles constitute structured representations of registered users. In many identification and authentication systems, such profiles are operationalized as machine learning–based classifiers that determine whether an incoming user matches a stored identity [9]. However, our framework treats the user profile not only as a classification reference, but also as a reusable embedding representation derived from biometric and trait-based data, shared across both authentication and behavior generation.

For biometric and behavioral data (e.g., head movement patterns, gesture dynamics, gaze trajectories, and optional physiological signals), we derive stable user-specific features that capture how a person typically moves and interacts. Rather than modeling raw signals directly, these patterns are summarized into a structured behavioral representation that reflects characteristic timing, intensity, and variability. In addition to functioning as a reference profile for authentication, this representation is used as supervision to fine-tune the nonverbal behavior generator (e.g., by updating a small set of personalization parameters) so that synthesized gestures and motion trajectories exhibit the same timing, variability, and expressive signatures present in the user's data.

For trait-based data (e.g., personality descriptions, autobiographical texts, prior conversation transcripts), we derive linguistic embeddings or structured descriptors using pretrained language models. These representations are incorporated into large language models through prompt conditioning, retrieval-augmented generation, or lightweight adapter-based fine-tuning. This enables verbal outputs that align with the registered user's tone, vocabulary, and communicative style.

3. Discussion

3.1. Research Prospects and Challenges

Mapping Authentication Signals to Perceived Identity Legitimacy. The key novelty of this conceptual framework lies in leveraging user profiles to personalize user-resembling agentic avatars in ways that support perceived identity legitimacy. Rather than introducing separate or highly individualized personalization models, our framework selectively utilizes identification cues already employed for authentication to inform avatar behavior. Importantly, no additional signals beyond those required for verification are introduced. By reusing authentication-relevant signals that encode stable and recognizable behavioral features, the framework links verified identity with obervable identity-consistent expression. Because these features reflect habitual interaction characteristics, they may influence how others attribute and recognize identity in mediated interaction. However, mapping authentication signals to perceived identity authenticity raises open questions. Features effective for identity classification are not necessarily perceptually salient in social interaction, and thus may not translate directly into recognizable behavioral cues. Determining how strongly authentication-derived patterns should shape behavior generation also requires calibration, as over-personalization may constrain behavioral flexibility. Empirical evaluation is therefore necessary to assess whether reusing authentication signals meaningfully enhances perceived identity continuity compared to conventional personalization strategies.

Designing Verifiable yet Privacy-Preserving User-Resembling Avatars. Because authentication-derived representations are reused to condition avatar behavior, the user profile becomes a shared resource across verification and generation. This dual use amplifies privacy risks if sensitive biometric or trait-based data are centrally stored or transmitted. A key research direction is therefore to perform feature extraction and identity matching locally on the user's device, exposing only minimal authentication outcomes to the application layer, consistent with prior arguments for embedding privacy protections directly into immersive system architectures [16]. Differential privacy techniques may further reduce re-identification risks while preserving user-specific structure [17]. Trait-based data introduce additional challenges: autobiographical text and stylistic samples used to condition language models may contain latent identifiers, and large language models have been shown to memorize and expose training data under adversarial probing [18]. Within our framework, this necessitates careful minimization of external data transmission, preference for on-device conditioning or retrieval mechanisms, and explicit constraints on how personal linguistic data are adapted or retained.

Designing Bias-Aware Profile-Derived Agents. Because user profiles directly condition verbal and nonverbal generation in our framework, any bias encoded in these representations may shape avatar outputs over time. Designing bias-aware profile-derived agents therefore requires treating the profile as an auditable design layer rather than a neutral data store. One approach is to distinguish stylistic identity markers (e.g., pacing, tone, expressiveness) from demographic or culturally loaded proxies, and to evaluate how behavior generation shifts when these dimensions are varied independently. Furthermore, because autobiographical traits may be incorporated into persistent conditioning signals, safeguards are necessary to prevent unintended reshaping of personal narratives, particularly given the reconstructive nature of human memory. Maintaining user-authored content as reviewable and versioned profile components can help preserve agency while enabling personalization.

3.2. Limitations and Future Work

This work presents a conceptual framework and does not include an implemented system or empirical validation. The feasibility of constructing stable shared profile representations across authentication and behavior generation remains to be demonstrated. In particular, it is unclear whether authentication-relevant features are consistently perceptually salient enough to influence perceived identity authenticity in real interaction contexts.

Future work will therefore focus on translating the framework into a functional prototype and empirically evaluating its effectiveness. We plan to implement a system that operationalizes the shared user profile representation across authentication and behavior generation pipelines. This includes (1) constructing biometric and trait-based profile embeddings during registration, (2) enabling on-device identity verification using incoming behavioral signals, and (3) integrating profile-conditioned mechanisms into nonverbal behavior synthesis and language generation modules.

We will conduct a controlled user study to evaluate whether reusing authentication-derived signals improves perceived identity legitimacy compared to baseline personalization approaches. Participants will interact with avatars under different personalization conditions in both synchronous and asynchronous scenarios. Dependent measures will include perceived identity legitimacy and continuity, social presence, and recognition accuracy. Through this evaluation, we aim to assess whether linking authentication-derived representations to avatar behaviors meaningfully supports identity-authentic avatar-mediated communication in Social XR.

4. Conclusion

In this position paper we proposed a conceptual framework that links authentication-derived signals to the perceived identity legitimacy of a registered, user-resembling avatar in Social XR. By reusing biometric and trait-based cues collected for identity verification as conditioning resources for behavior

generation, the framework integrates authentication and personalization through a shared user profile representation. Rather than separating verification from expressive behavior, this approach seeks to support identity continuity across synchronous and asynchronous interaction. We further discuss key research prospects and challenges to guide future implementation and empirical validation of this approach.

Acknowledgments

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [RGPIN-2022-03294].

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT in order to: Grammar and spelling check. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] J. N. Bailenson, J. Blascovich, A. C. Beall, J. M. Loomis, Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments, Presence: Teleoperators and Virtual Environments 10 (2001) 583–598. URL: https://direct.mit.edu/pvar/article/10/6/583-598/18392. doi:10.1162/105474601753272844.

[2] F. Herrera, S. Y. Oh, J. N. Bailenson, Effect of Behavioral Realism on Social Interactions Inside Collaborative Virtual Environments, PRESENCE: Virtual and Augmented Reality 27 (2018) 163–182. URL: https://direct.mit.edu/pvar/article/27/2/163/96071/Effect-of-Behavioral-Realism-on-Social. doi:10.1162/pres_a_00324.

[3] S. Baker, J. Waycott, R. Carrasco, R. M. Kelly, A. J. Jones, J. Lilley, B. Dow, F. Batchelor, T. Hoang, F. Vetere, Avatar-Mediated Communication in Social VR: An In-depth Exploration of Older Adult Interaction in an Emerging Communication Platform, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 2021, pp. 1–13. URL: https://dl.acm.org/doi/10.1145/3411764.3445752. doi:10.1145/3411764.3445752.

[4] J. McVeigh-Schultz, K. Isbister, A "beyond being there" for VR meetings: envisioning the future of remote work, Human–Computer Interaction 37 (2022) 433–453. URL: https://www.tandfonline.com/doi/full/10.1080/07370024.2021.1994860. doi:10.1080/07370024.2021.1994860.

[5] A. Giovannelli, L. Pavanatto, S. Davari, H. Miao, V. Chheang, B. Giera, T. Bremer, D. A. Bowman, Investigating the influence of playback interactivity during guided tours for asynchronous collaboration in virtual reality, in: 2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR), IEEE, Saint Malo, France, 2025, p. 23–33. URL: https://ieeexplore.ieee.org/document/10937462/. doi:10.1109/VR59515.2025.00027.

[6] P. Wang, M. R. Miller, A. C. Queiroz, J. N. Bailenson, Socially late, virtually present: The effects of transforming asynchronous social interactions in virtual reality, in: Proceedings of the CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA, 2024, p. 1–19. URL: https://dl.acm.org/doi/10.1145/3613904.3642244. doi:10.1145/3613904.3642244.

[7] H. Nam, M. Sarvesh, S. Kang, W. Woo, K. Kim, Effects of ai-powered embodied avatars on communication quality and social connection in asynchronous virtual meetings, IEEE Transactions on Visualization and Computer Graphics 31 (2025) 10152–10162. doi:10.1109/TVCG.2025.3616761.

[8] J. Lin, M. E. Latoschik, Digital body, identity and privacy in social virtual reality: A systematic review, Frontiers in Virtual Reality 3 (2022) 974652. URL: https://www.frontiersin.org/articles/10.3389/frvir.2022.974652/full. doi:10.3389/frvir.2022.974652.

[9] I. Anastasaki, G. Drosatos, G. Pavlidis, K. Rantos, User Authentication Mechanisms Based on Immersive Technologies: A Systematic Review, Information 14 (2023) 538. URL: https://www.mdpi.com/2078-2489/14/10/538. doi:10.3390/info14100538.

[10] F. Chiossi, C. Ou, S. Mayer, Optimizing Visual Complexity for Physiologically-Adaptive VR Systems: Evaluating a Multimodal Dataset using EDA, ECG and EEG Features, in: Proceedings of the 2024 International Conference on Advanced Visual Interfaces, AVI '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: https://doi.org/10.1145/3656650.3656657. doi:10.1145/3656650.3656657, event-place: Arenzano, Genoa, Italy.

[11] K. Pfeuffer, M. J. Geiger, S. Prange, L. Mecke, D. Buschek, F. Alt, Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland Uk, 2019, pp. 1–12. URL: https://dl.acm.org/doi/10.1145/3290605.3300340. doi:10.1145/3290605.3300340.

[12] J. Liebers, M. Abdelaziz, L. Mecke, A. Saad, J. Auda, U. Gruenefeld, F. Alt, S. Schneegass, Understanding User Identification in Virtual Reality Through Behavioral Biometrics and the Effect of Body Normalization, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, Yokohama Japan, 2021, pp. 1–11. URL: https://dl.acm.org/doi/10.1145/3411764.3445528. doi:10.1145/3411764.3445528.

[13] J. Liebers, P. Laskowski, F. Rademaker, L. Sabel, J. Hoppen, U. Gruenefeld, S. Schneegass, Kinetic Signatures: A Systematic Investigation of Movement-Based User Identification in Virtual Reality, in: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, CHI '24, Association for Computing Machinery, New York, NY, USA, 2024. URL: https://doi.org/10.1145/3613904.3642471. doi:10.1145/3613904.3642471, event-place: Honolulu, HI, USA.

[14] P. Papadopoulou, A. Andreou, P. Kanellis, D. Martakos, Enabling trust formation within agent-mediated virtual environments, Virtual Reality 5 (2000) 149–159. URL: http://link.springer.com/10.1007/BF01409420. doi:10.1007/BF01409420.

[15] D. H. Mcknight, M. Carter, J. B. Thatcher, P. F. Clay, Trust in a specific technology: An investigation of its components and measures, ACM Transactions on Management Information Systems 2 (2011) 1–25. URL: https://dl.acm.org/doi/10.1145/1985347.1985353. doi:10.1145/1985347.1985353.

[16] F. Roesner, T. Kohno, D. Molnar, Security and privacy for augmented reality systems, Communications of the ACM 57 (2014) 88–96. doi:10.1145/2580723.2580730.

[17] C. Dwork, Differential Privacy, volume 4052 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, p. 1–12. URL: http://link.springer.com/10.1007/11787006_1. doi:10.1007/11787006_1.

[18] N. Carlini, F. Tramèr, E. Wallace, M. Jagielski, A. Herbert-Voss, K. Lee, A. Roberts, T. B. Brown, D. Song, Ú. Erlingsson, A. Oprea, C. Raffel, Extracting training data from large language models, CoRR abs/2012.07805 (2020). URL: https://arxiv.org/abs/2012.07805. arXiv:2012.07805.