

Latent Augmentation: AI-Enabled XR and the Erosion of Unwatched Social Interaction

Anonymous Authors
Blind Submission for Review

Abstract—Wearable extended reality (XR) embeds computation directly into everyday physical and social encounters. Artificial intelligence extends the capacity of such devices from merely recording images to continually identifying faces, linking observed subjects to external databases, indexing encounters into searchable memories, and modifying the wearer’s perceptual experience. Critically, these processes happen without the conscious inspection by either party of the data that went into the computation. We term this process latent augmentation. This paper provides a structured conceptual framework built on five key capabilities: identification, inference, memory, real-time assistance, and adaptive perception. The five capabilities can create capability asymmetries between co-present individuals who appear to be in the same physical situation but have different access to machine-processed data. Their social effects differ depending on whether the affected person is aware, unaware, or uncertain about their application. While it is recognized in scholarship that smart glasses bring general surveillance risks and information asymmetries between the observer and the bystander, the innovation here is the structured five-capabilities by three-awareness-states framework to understand how latent and non-visual AI interpretation creates specific vulnerabilities. Finally, we propose that regulation and system design must consider downstream machine inferences as the critical boundary point and not simply sensor-based data collection upfront.

Index Terms—Extended reality, smart glasses, latent augmentation, capability asymmetry, bystander privacy, inference governance, practical obscurity.

I. INTRODUCTION

The abilities enabled through human technology have always expanded the set of possible actions available to a person. A match makes the creation of fire possible on demand, a flashlight makes vision possible in the dark, and a camera makes an image available which would otherwise vanish along with the event itself. Technology expands the range of possible actions, yet their results can generally be accessed consciously by the user of the technology. It is possible to see the image in a photograph, hear a recording, or access information stored on a device. Even though the action may be facilitated through technology, the result can still be perceived consciously by the user.

In terms of artificial intelligence, the relation between a human and a result changes because the useful output does not necessarily have to be examined consciously. A camera may take a picture while an AI system recognizes the person in the picture, retrieves information about that person from its memory, compares the current encounter with previous ones, and makes an inference without the user having to look at the original picture [1]–[3]. The same process can be performed

with speech, eye movements, physical movements, the location of a person, or other data which may be collected by an XR device [4]–[6]. In this case, the informative output is not what the device captures, but what the system is capable of deducing from the capture and how that influences the situation afterward. An observation can become an inference, and an inference can influence a person’s actions without anyone knowing what happened in the middle.

Wearable XR creates a situation where this relation is placed directly into physical interaction. Two people may be in the same space, see the same things, and appear to be communicating with each other, yet they may perceive completely different sets of information regarding that interaction. While one may see the face in front of them, the other may have additional capabilities such as recognizing the identity of the person, remembering encounters they had previously, inferring various attributes of that person, accessing context-specific information, or receiving instructions on how to act during the interaction [7], [8]. The difference may not be noticeable during the interaction, thus creating an informational asymmetry which is not apparent from the co-presence of the people involved.

The effect also depends on whether the person understands that there is an informational asymmetry. If a person knows that they may be recognized, interpreted, or permanently remembered, they may change their behavior even before an inference is made [9]. However, if a person is unaware of the process, they may be influenced by decisions made by an AI based on its interpretation of the interaction. Between these conditions is uncertainty about the potential capabilities of the device. Even if smart glasses are not recording something at the moment, another person may consider whether they are [10]. Therefore, the possibility of observation can have an impact even if the capability is not activated, while an activated capability may influence the interaction without being revealed.

This paper makes three contributions. First, we define latent augmentation as machine-processed XR output whose intermediate processing is not fully visible to the affected person. Second, we cross five capabilities (identification, inference, memory, real-time assistance, adaptive perception) with aware, unaware, and uncertain states. Third, we identify design and governance implications for downstream machine inferences in social interaction.

FIVE CAPABILITIES × THREE AWARENESS STATES

Affected person's awareness of capability application			
CAPABILITY	AWARE knows it is active	UNAWARE does not know	UNCERTAIN cannot verify
1 Identification	May limit disclosure	Identity linked unseen	May guard identity cues
2 Inference	May manage visible cues	Attributes inferred unseen	May monitor own cues
3 Memory	Expects later recall	Encounter saved unseen	May question retention
4 Real-time assistance	May read reply as aided	Prompt shapes reply unseen	May question spontaneity
5 Adaptive perception	Accounts for altered view	Altered view unseen	May doubt shared view

Fig. 1. The Latent Augmentation Framework. Sensor streams in multimodal XR are autonomously translated into five latent capabilities, resulting in latent capability disparities, which produce different impacts socially depending on whether the relevant parties are informed, uninformed, or unsure.

II. RELATED WORK AND CONCEPTUAL GROUNDING

Current XR systems already capture types of data that can reveal information beyond what is necessary for the interaction itself. For instance, eye-tracking data not only captures attention but may also expose private attributes of the person depending on information about the stimulus they are looking at [5], [11]. It has been shown that gaze patterns can be used for modeling decision-making and other aspects of the person, thus creating the potential for a system to predict and perhaps even influence the decisions of the user [12]–[14]. As such, the privacy issues associated with sensor data collection do not end with the collection itself, because additional information may be extracted from what the sensor observed even though it was never provided explicitly by the user.

Another aspect of the societal impact of wearable XR that has been explored concerns the people around the wearer. According to the findings of Chung et al., there is a power imbalance in interactions between a person wearing AR glasses and others [7]. The inequality is caused by the fact that the person not wearing any additional equipment cannot see the information that the wearer sees or control the way they are represented. This informational disadvantage exists even before any kind of AI is introduced. Moreover, the question of modifying perception has also been studied in AR systems used in real-life settings [15]–[17].

The introduction of AI into the equation makes the possible outputs of the system even broader. For instance, SocialMind captures information from a live conversation using multimodal sensing and then generates recommendations for the person wearing AR glasses [8]. AR Secretary Agent can also identify the interlocutor and extract information based on previous interactions, making it possible for the wearer to access information that he would have had to remember himself [3]. From this perspective, the technology is doing some of the job of remembering for the user. This setup represents cognitive offloading, where some functions of memory and recall are transferred from the person to the system [18], [19]. It is

evident from the above examples that the wearer need not do each act of identification, inference, memory recall, or help separately as individual mental acts. The wearer can just switch on the system, and the intervening process will be taken care of by the system itself.

Prior research has examined these capabilities in separate contexts. Work on wearable computing considers bystander privacy and awareness [7], [10]; other work examines inference from XR sensor data [4]–[6], augmented memory [2], [20], live assistance [8], and perceptual modification [15], [16]. We bring these strands together by crossing a system’s downstream capability with the awareness of the affected person, so that the same capability can be examined across aware, unaware, and uncertain encounters.

To construct the framework, we considered work on XR sensor privacy and inference, informational asymmetry in wearable technology, augmented memory and assistance, and perceptual modification. We grouped recurring functions by what the system can do during an interaction, distinguishing identification, inference, memory, real-time assistance, and adaptive perception. We separated awareness into aware, unaware, and uncertain states because the social impact of a capability may depend on whether the affected person recognizes its use. This is a targeted conceptual synthesis rather than a systematic review of the XR literature.

We refer to this situation as latent augmentation and define it as a case where collected information is processed into identification, inference, memory, assistance, or perceptual adaptation without the intermediate process being fully visible to the affected person.

III. LATENT AUGMENTATION FRAMEWORK

For our purposes, we distinguish the following kinds of capabilities used for augmentation: identification, inference, memory, real-time assistance, and adaptive perception (Fig. 1). By themselves, these capabilities may not be inherently harmful because, in many cases, they might help to increase the utility of the system for its user. Problems arise when one

of these capabilities affects what a person is able to see or understand during an interaction while the second person does not have access to the same capability or know that it is being applied to them. Hence, these capabilities should be viewed as possible causes of asymmetry.

First, the capability of identification means that the system can identify a person and connect their identity to external information [1]. Second, the capability of inference is similar to identification in the sense that it allows information to be derived from sensor data. However, while identification connects directly observable features to an existing identity, inference may derive information which was not directly provided by the person. Namely, a person's attention, emotional state, preferences, and intentions may be derived from sensor data [5], [6], [12]. Third, the capability of memory allows previous interactions to be saved and made available for future reuse [2], [20]. Fourth, the capability of real-time assistance involves providing certain recommendations or pieces of information to the user while interacting with another person [8]. Fifth, the capability of adaptive perception involves changing what is visible, audible, or otherwise perceptible to the user [1], [15], [16]. Despite the different purposes of these capabilities, each of them changes the conditions under which the interaction occurs.

The impact of the above-listed capabilities is heavily dependent on the person's awareness regarding their use. We distinguish between three kinds of awareness: aware, unaware, and uncertain. Aware means that the person knows that some form of augmentation is taking place and, hence, can change their behavior because of this. Unaware means that a person does not know that some capability is being used on them and, hence, may behave in the usual way while still being affected by its consequences. Uncertain means that the person knows that the device may have these capabilities but does not know whether they are being applied. It is possible to be aware of a capability being active without knowing about what data it works with and how it generates the output. Three states of awareness pertain to awareness of capability application.

The combination of capability and awareness results in different kinds of asymmetries. Thus, for example, identification may result in a person limiting what they reveal when they are aware of this capability, while another person may remain unaware that their information has already been retrieved. Memory may result in a difference between people who are aware that their encounter will be saved in some form of memory and those who are unaware of this fact and, hence, expect it to disappear into normal human forgetfulness [20], [21]. Finally, real-time assistance and adaptive perception may be even harder to notice for the affected person because they may not know that additional information is being received or that a modification of perception is occurring.

IV. SOCIAL EFFECTS OF AWARENESS

This is how the technical aspects of asymmetry should be described. As for the social ones, they are reflected in the responses of the people involved in the interaction. When a per-

son realizes that they can become an object of identification, inference, and memory operations, it does not really matter whether any technology was actually applied. Just thinking about its possible use can change people's behavior [9], [10]. In other words, the result of the augmentation may take place before the actual augmentation itself. In this way, the person can decide not to give away any particular piece of information, not to do anything specific, and to behave in a certain way which will look socially appropriate considering possible future inferences.

In the same way, an uncertain situation can affect behavior since the person does not know which exact capabilities of the device are currently working [10]. For instance, while interacting with someone who wears smart glasses, the person may know that the device is capable of recording, identification, information retrieval, or providing other forms of assistance, yet there is no way to find out which exact capabilities are working at a given moment. As a consequence, the interaction is influenced not only by the currently active capabilities of the device but by all of its possible capabilities as well. Uncertainty becomes a factor in the interaction and changes people's behavior similarly to known observation.

Another problem caused by unawareness is that the affected person is unable to react to the asymmetry. When the system performs identification of the person, makes some inferences regarding them, or gives a suggestion to the user without informing them about this process, the resulting behavior of the user may appear to be absolutely spontaneous and free from any external influence. Thus, the AI affects the interaction indirectly, while the person reacts to the decision, recommendation, or change of behavior without knowing that machine-processed information played a role in causing it.

As a conclusion, it should be stated that the problem is much broader than privacy. Information may be unknown to both participants, but it can still have consequences for one of them. The general problem is that a person's ability to interact in a common space without the constant transformation of their actions into machine-readable information, which may later have some impact on this interaction or another one, is reduced. Privacy is a part of the general issue, yet the core of the problem is whether social interactions can remain temporary and ambiguous [21]–[23].

V. EMERGENT INCENTIVES AND GOVERNANCE

Such asymmetry in the relationship between the parties involved in the development of surveillance technologies does not require any special intentions from these parties. The user of XR may only want to get more information, while the platform will strive to increase the utility of the device, providers will strive to increase the adoption of their systems, and institutions will try to gain an advantage from available information. There is nothing contradictory in this case, and these needs may complement each other in the following manner: more sensing can enable more inferences and increase utility, while more adoption can bring more information needed for system improvement [24], [25]. Thus,

such a condition may emerge even without anyone intending it.

It means that we deal with the problem of positive feedback here because the usefulness of latent augmentation gives a motive to increase its extent. If identification gives an easier way to interact, there is a motive to improve identification. If past interactions can be remembered, there is a motive to remember more of them. If assistance can be personalized using some information about a person, there is a motive to increase sensing and thus increase the specificity of assistance. Thus, different capabilities, initially aimed at solving different problems, may gradually connect in a single system where sensing, inference, memory, and assistance occur continuously.

It also means that regulation aimed exclusively at limiting the amount of sensor data can be ineffective for the following reason. The same data can have very different effects depending on how it is used and what inference can be made from it. Eye movement data making it possible to improve image rendering is quite different from eye movement data used to draw an inference about the uncertainty of a person and thus change their advertisements or provide an opportunity to influence them in some way. Thus, the problem is not only what data can be collected but also what inference can be drawn from this data and how it can be used [11], [14], [26], [27].

The principle that may be used for the regulation of latent augmentation is the visibility of consequential capability use to the people affected by it [17], [27]. Of course, this cannot mean the disclosure of all computations made within the system because it may make many systems ineffective. However, it means that a person affected by the system should be aware that they are interacting with a system having some capability of identification, inference, memory, assistance, or adaptive perception, especially when this capability affects them. The purpose of this limitation is not the prohibition of augmentation but the prevention of useful capabilities from becoming invisible tools for influencing a person without providing them with an opportunity to understand and respond to this asymmetry.

VI. DISCUSSION AND FUTURE WORK

The purpose of the framework is to offer a way to characterize different types of latent augmentation and their influence on the nature of the interaction. The five capabilities can occur either individually or in combination, while the level of awareness on the part of the affected person can vary during the process of interaction. Thus, it becomes possible not only to characterize the technical capabilities of the system but also to demonstrate that the same capability may have different social consequences depending on the level of awareness of its use.

It will be helpful to examine which combinations of capabilities can be observed in particular XR environments and which of them contribute the most to asymmetry. Also, it might be useful to investigate which levels of disclosure are sufficient to reduce the consequences of uncertainty without

complicating the system's functionality. Thus, one can say that the framework is rather an attempt at structuring further research into the social aspects of the latent augmentation phenomenon.

VII. CONCLUSION

It is the development of wearable XR that enables the involvement of artificial intelligence in physical and social interactions. Nevertheless, the most significant and unique aspect of this type of interaction is that information can have an impact on this interaction without anyone consciously analyzing the initial information and the process of its interpretation. Depending on the type of interaction and the purpose being served, the information obtained during the interaction can be utilized for identification, inference, memory, real-time assistance, and adaptive perception, hence changing the interaction without revealing the intermediate process. This phenomenon is termed latent augmentation, which is described in terms of five capabilities and three states of awareness.

The social issue raised by the phenomenon of latent augmentation goes beyond the scope of privacy infringement. In case a person knows that he or she might be identified, analyzed, or remembered by a machine, he or she may change their behavior before the capability is actually employed. The absence of awareness allows these capabilities to affect the interaction without allowing the person to respond. The outcome of this process will be an asymmetry in the interaction due to the fact that the technical conditions remain hidden behind the shared physical space.

In this way, the integration of wearable XR into people's lives will make it necessary to think about the capabilities of these machines not just in terms of the information that they can sense, but rather in terms of what kinds of machine-processed information are permissible to influence interactions between people. However, the preservation of the usefulness of augmentation does not mean that the consequences of its employment should remain invisible to people.

REFERENCES

- [1] F. Roesner, T. Kohno, and D. Molnar, "Security and privacy for augmented reality systems," *Commun. ACM*, vol. 57, no. 4, pp. 88–96, 2014.
- [2] B. J. Rhodes, "The wearable remembrance agent: A system for augmented memory," *Personal Technologies*, vol. 1, no. 4, pp. 218–224, 1997.
- [3] R. A. El Haddad, Z. Wang, Y. Shin, R. Liu, Y. Wang, and C. Yu, "AR Secretary Agent: Real-time memory augmentation via LLM-powered augmented reality glasses," arXiv:2505.11888, 2025. [Online]. Available: <https://arxiv.org/abs/2505.11888>
- [4] G. Munilla Garrido, V. Nair, and D. Song, "SoK: Data privacy in virtual reality," *Proc. Privacy Enhancing Technol.*, vol. 2024, no. 1, pp. 21–40, 2024.
- [5] J. L. Kröger, O. H.-M. Lutz, and F. Müller, "What does your gaze reveal about you? on the privacy implications of eye tracking," in *Privacy and Identity Management. Data for Better Living: AI and Privacy*, ser. IFIP Adv. Inf. Commun. Technol. Springer, 2020, vol. 576, pp. 226–241.
- [6] V. Nair, C. Rack, W. Guo, R. Wang, S. Li, B. Huang, A. Cull, J. F. O'Brien, M. Latoschik, L. Rosenberg, and D. Song, "Inferring private personal attributes of virtual reality users from ecologically valid head and hand motion data," in *Proc. IEEE Conf. Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2024, pp. 477–484.

- [7] J. W. Chung, J. Fu, Z. Deocadiz-Smith, M. F. Jung, and J. Huang, "Negotiating dyadic interactions through the lens of augmented reality glasses," in *Proc. ACM Designing Interactive Systems Conf. (DIS)*, 2023, pp. 493–508.
- [8] B. Yang, Y. Guo, L. Xu, Z. Yan, H. Chen, G. Xing, and X. Jiang, "SocialMind: LLM-based proactive AR social assistive system with human-like perception for in-situ live interactions," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 9, no. 1, pp. 23:1–23:30, 2025.
- [9] J. W. Penney, "Chilling effects: Online surveillance and Wikipedia use," *Berkeley Technology Law Journal*, vol. 31, no. 1, pp. 117–182, 2016.
- [10] T. Denning, Z. Dehlawi, and T. Kohno, "In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies," in *Proc. SIGCHI Conf. Human Factors in Computing Systems (CHI)*, 2014, pp. 2377–2386.
- [11] S. A. Byrne, N. Castner, E. Bozkir, D. C. Niehorster, and E. Kasneci, "From lenses to living rooms: A policy brief on eye tracking in XR before the impending boom," in *Proc. IEEE Int. Conf. Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, 2024, pp. 90–96.
- [12] B. David-John, C. E. Peacock, T. Zhang, T. S. Murdison, H. Benko, and T. R. Jonker, "Towards gaze-based prediction of the intent to interact in virtual reality," in *Proc. ACM Symp. Eye Tracking Research and Applications (ETRA Short Papers)*, 2021, pp. 2:1–2:7.
- [13] K. Yeung, "hypernudge": Big data as a mode of regulation by design," *Information, Communication & Society*, vol. 20, no. 1, pp. 118–136, 2017.
- [14] R. Calo, "Digital market manipulation," *George Washington Law Review*, vol. 82, no. 4, pp. 995–1051, 2014.
- [15] S. Mann and J. Fung, "EyeTap devices for augmented, deliberately diminished, or otherwise altered visual perception of rigid planar patches of real-world scenes," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 2, pp. 158–175, 2002.
- [16] S. Mori, S. Ikeda, and H. Saito, "A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects," *IPSJ Trans. Comput. Vis. Appl.*, vol. 9, no. 1, p. 17, 2017.
- [17] J. O'Hagan, J. Gugenheimer, F. Mathis, J. Bonner, R. Jones, and M. McGill, "A viewpoint on the societal impact of everyday augmented reality and the need for perceptual human rights," *IEEE Security & Privacy*, vol. 22, no. 1, pp. 64–68, 2024.
- [18] E. F. Risko and S. J. Gilbert, "Cognitive offloading," *Trends Cogn. Sci.*, vol. 20, no. 9, pp. 676–688, 2016.
- [19] B. Sparrow, J. Liu, and D. M. Wegner, "Google effects on memory: Cognitive consequences of having information at our fingertips," *Science*, vol. 333, no. 6043, pp. 776–778, 2011.
- [20] N. Davies, A. Friday, S. Clinch, C. Sas, M. Langheinrich, G. Ward, and A. Schmidt, "Security and privacy implications of pervasive memory augmentation," *IEEE Pervasive Computing*, vol. 14, no. 1, pp. 44–53, 2015.
- [21] V. Mayer-Schönberger, *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press, 2009.
- [22] W. Hartzog and F. Stutzman, "The case for online obscurity," *California Law Review*, vol. 101, no. 1, pp. 1–49, 2013.
- [23] J. E. Cohen, "What privacy is for," *Harvard Law Review*, vol. 126, no. 7, pp. 1904–1933, 2013.
- [24] P. E. Agre, "Surveillance and capture: Two models of privacy," *The Information Society*, vol. 10, no. 2, pp. 101–127, 1994.
- [25] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs, 2019.
- [26] H. Nissenbaum, "Privacy as contextual integrity," *Washington Law Review*, vol. 79, no. 1, pp. 119–158, 2004.
- [27] S. Wachter and B. Mittelstadt, "A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI," *Columbia Business Law Review*, vol. 2019, no. 2, pp. 494–620, 2019.