The Lens Is Not the Problem: Understanding Potential Privacy Risks and User Concerns in AI Glasses

Mohammed Aldeen
Clemson University
South Carolina, USA

Jingwen Yan
Clemson University
South Carolina, USA

Zixun Xiong
George Mason University
Virginia, USA

Yicheng Zhang
George Mason University
Virginia, USA

Long Cheng
Clemson University
South Carolina, USA

Abstract

AI glasses are emerging as one of the fastest-growing categories in consumer wearables, embedding always-on cameras, microphones, and cloud-connected multimodal assistants. Yet these same sensors can continuously capture audio, video, and biometric signals from both wearers and nearby people, routing much of that data to cloud servers and human reviewers through a pipeline that can expose intimate scenes, faces, and conversations without consent. In this paper, we analyze the privacy risks of Meta's AI glasses through two complementary lenses: an audit of the company's published documentation and a large-scale study of how users discuss these glasses online. First, we organize Meta's privacy policies and data practices into six data channels that trace how captured signals move from the device through cloud and we use the LINDDUN privacy framework to classify the resulting threats across wearers, bystanders, and external adversaries. We then analyze 4,900 public discussions from Reddit about users' privacy concerns of Meta's AI glasses, and we find that public concern fixates on the visible act of recording while rarely naming that captured data is stored in the cloud, used to train AI models, or reviewed by humans. The privacy story of AI glasses, we argue, is less about the lens than about everything that happens after capture.

CCS Concepts

• **Security and privacy → Social aspects of security and privacy**.

ACM Reference Format:
Mohammed Aldeen, Jingwen Yan, Zixun Xiong, Yicheng Zhang, and Long Cheng. 2026. The Lens Is Not the Problem: Understanding Potential Privacy Risks and User Concerns in AI Glasses. In *Second Workshop on Enhancing Security, Privacy, and Trust in Extended Reality (XR Security '26), October 26–30, 2026, Austin, TX, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3842203.3844596

This work is licensed under a Creative Commons Attribution 4.0 International License.
XR Security '26, Austin, TX, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2951-5/26/10
https://doi.org/10.1145/3842203.3844596

1 Introduction

Smart glasses have quickly moved from experimental prototypes into mainstream consumer products. A prominent example is the Ray-Ban Meta series, which has surpassed two million units sold since launch [12]. Unlike VR headsets, which users put on for bounded sessions in controlled environments, AI glasses accompany their wearers into homes, workplaces, classrooms, hospitals, stores, and public streets, where they capture both the wearer's private activities and the surroundings of nearby bystanders [26, 27]. This everyday deployment model makes the privacy implications more pervasive, as prior studies have analyzed the privacy implications of smart glasses, as discussed further in Appendix A. At the same time, AI glasses are increasingly connected to cloud-based LLMs and VLMs through a paired phone or wireless network, turning whatever the wearer sees and hears into an online AI query [18, 20]. As a result, privacy risks are no longer limited to local camera or microphone access; they also extend to the network traffic generated when users ask AI models about what they see, hear, or do in the physical world.

Meta AI glasses are multimodal wearable devices that can collect and process diverse information from different modalities around the wearer. The camera modality can capture photos, videos, faces, documents, screens, places, and bystanders. The smart speaker modality can capture spoken commands, ambient speech, background sound, and other sounds [21]. Meta's AI features can also link these modalities together into a continuous, live stream. For example, Live AI keeps the camera and microphone actively running at the same time during a session, allowing the assistant to look, listen, and respond to the wearer's surroundings in real time [23]. Also, the Autocapture feature uses contextual understanding to decide which moments should be captured,

and depends on Meta's cloud to process the captured media [22]. Together, privacy risks in these glasses develop through a sequence: the glasses first sense the physical environment, then convert what they capture into digital audio, image, video, transcript, metadata, then moved across the Meta AI mobile app, Meta's cloud services, persistent storage, and the machine-learning and human-review processes used to improve Meta products [21]. As a result, simply being near someone wearing the glasses can become a source of data. A normal moment in physical space can be turned into information that software can be stored, searched, analyzed, or train models without any consent. We provide an extended discussion of related work in Appendix A.

To make this sensing pipeline more visible and controllable, Meta provides privacy notices and controls at several layers. At the device level, Ray-Ban Meta glasses use a visible capture LED to indicate photo, video, or livestream capture, and the device notifies the wearer if the LED is covered [19]. At the application level, the Meta AI mobile app provides privacy settings, permission prompts, and access to product-specific privacy notices [24]. Meta also discloses app-level data practices through platform mechanisms such as Apple's App Privacy labels, Google Play's Data Safety section, and regional privacy notices [4, 15]. This creates a gap between device behavior and user understanding: a bystander may see eyewear or an LED without knowing whether the glasses are idle, listening for voice activation, recording, using Live AI, or sending data to the cloud. Similarly, a wearer may see privacy settings without fully understanding how captured data can later be retained, reviewed, or reused. This gap motivates our study of how users discuss and interpret privacy concerns around Meta AI glasses.

To close this gap, this position paper analyzes Meta AI glasses from two complementary perspectives. First, we analyze Meta's official privacy notices and product documentation, extract the device behaviors that Meta itself describes, and trace how captured data move through the device, the paired app, and Meta's cloud. Second, we examine how users discuss these risks in practice by analyzing Reddit comments about Meta AI glasses and identifying the concerns that users repeatedly raise about recording, data processing, bystander exposure, and Meta's trustworthiness. Therefore, this paper makes three contributions:

- We identify privacy-relevant behaviors from Meta's official documentation and organize them into a six-channel pipeline that shows how data enters the glasses, moves through the paired app and cloud services, and may later reach storage, review, or reuse.
- We define three threat models that distinguish platform-side risks to the wearer, bystander exposure, and external-adversary risks. We then map these risks to the LINDDUN privacy framework to clarify who is harmed, where the harm originates, and which privacy categories each risk triggers.
- We conduct large-scale analysis on Reddit discussions to show how users frame the privacy risks of Meta AI glasses in practice.

2 Privacy Risks of Meta AI Glasses

We begin by examining what Meta's own documentation reveals about the glasses' sensing capabilities and data practices, tracing how captured information moves through the device, the paired app, and Meta's cloud infrastructure. To do that, we explore how privacy risks develop in four stages. First, we establish a factual baseline using Meta's own documentation to analyze privacy risks (Section 2.1). Second, we breakdown Meta AI glasses device into six data channels to show how information is transferred from the physical world to become digital data that can be processed, stored, or analyzed (Section 2.2). We then categorize the risks into three distinct threat models: risks to the user, risks to bystanders, and risks created by external misuse or attack (Section 2.3). This structure makes clear who is affected, how the harm arises, and whether the privacy risk comes from the platform's ordinary design or from an outside actor. Finally, we apply the LINDDUN privacy framework to translate defined privacy threats into seven categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure, Unawareness, and Non-compliance (Section 2.4).

2.1 Privacy-Relevant Behaviors of Meta AI Glasses

To analyze privacy risks, we first need to understand how these AI glasses operates. We therefore analyze Meta's official privacy notices and product documentation to establish how the device captures, transmits, and processes data before we draw any conclusions about risk. To do that we systematically collected the relevant privacy policies, terms of service, and technical support pages directly from Meta's corporate website. We then manually reviewed these preserved pages to extract Meta's explicit claims regarding data practices and cloud processing. This extraction process yielded five core device behaviors that transform the glasses, as shown in Table 1. Their documentation discloses that these five behaviors are not independent features but sequential stages of a single pipeline, where the output of each stage becomes the input of the next. The glasses listen at all times for a wake word (B1), which means audio from the surrounding environment enters the system before the user ever speaks a command. Because accidental activations occur, as Meta itself acknowledges, ambient audio is also captured, uploaded, and retained in cloud storage for up to one year (B2). Once stored, those

recordings and transcripts become available to Meta's machine learning systems and to trained human reviewers (B3). Live AI then extends this chain further by activating the camera and microphone simultaneously throughout a session (B4), so the device no longer captures individual moments but streams the user's entire environment continuously. Autocapture extends it further still by deciding autonomously which moments to record (B5), removing even the requirement that the user initiate capture. The significance of this sequence is that a person simply wearing these glasses in a room causes the conversations, faces, and activities of everyone present to move through a pipeline that ends at cloud storage and human review, regardless of whether the wearer intended to record anything at all.

Table 1: Privacy-relevant device behaviors extracted from Meta's official documentation.

ID	Documented device behavior	Source basis
B1	Meta AI can be activated by saying "Hey Meta" or by pressing and holding the touchpad. Wake-word activation means the system must be able to detect the activation phrase before a user gives a command.	Meta AI glasses help pages; Meta voice privacy notice
B2	Meta defines voice interactions to include inadvertent activations and background sound that occurs during a voice interaction. Meta also states that it stores voice recordings even when the interaction was unintentionally activated. False wakes are deleted within 90 days of detection, while other stored voice recordings and transcripts may be retained for up to one year.	Meta voice privacy notice
B3	Stored voice recordings and transcripts may be processed using machine learning and trained human reviewers to improve Meta products and voice services.	Meta voice privacy notice
B4	During a Live AI session, the glasses' camera and microphone are continuously active so that Meta AI can respond to what the user is seeing and hearing.	Meta Live AI documentation
B5	Autocapture uses contextual understanding to decide what moments to capture. During an Autocapture session, Meta AI captures a continuous low-resolution series of images, records short video clips, and requires cloud media processing to be enabled. Meta also describes safeguards intended to prevent recording in sensitive locations such as bathrooms or locker rooms.	Meta Autocapture documentation

2.2 The Data Pipeline of Meta AI Glasses

To understand how the device behaviors in Section 2.1 create privacy risks, we need to trace where the captured information goes. We use the term *data channel* to denote a specific point in the device's data pipeline at which information is either acquired from the physical environment or moved between the glasses, the paired phone, and Meta's cloud. As shown in Table 2, we identify six channels, organized into two phases. The first phase covers acquisition, meaning the three ways data enters the device: through ambient listening for the wake word (C1), through photos and videos the user intentionally records (C2), and through continuous camera and microphone activity during Live AI and Autocapture sessions (C3). The remaining three (C4–C6) describe *where data goes* after capture: transfer off the device, cloud processing and storage, and review or reuse. A single Live AI clip, for example, enters at C3, transfers at C4, is processed and stored at C5, and may reach a human reviewer at C6. This sequential structure matters because privacy risk does not remain constant across the pipeline. At C1 and C2, the risk is primarily one of awareness: the wearer controls the device, but bystanders may not know they are being captured. The further the data travels through this pipeline, the further it moves from any meaningful control that either the wearer or the bystander can exercise over it.

Table 2: Data channels in Meta AI glasses, grouped by phase. C1–C3 describe how data is acquired from the environment; C4–C6 describe how acquired data moves through the platform.

Channel	Name	What flows through it	Evidence
Acquisition — how data enters the device?			
C1	Ambient sensing	Audio captured while the glasses listen for the wake word, including background sound and inadvertent activations.	B1, B2
C2	Explicit capture	Photos and videos the user intentionally takes via the capture button or a voice command.	–
C3	Session capture	Continuous camera and microphone during Live AI; low-resolution image streams and short clips during Autocapture.	B4, B5
Movement — where acquired data goes?			
C4	Transfer	Data moving from the glasses to the Meta AI mobile app, the phone's photo library, or Meta's cloud services.	B4, B5
C5	Cloud processing and storage	Cloud-side ASR, vision, and contextual inference; retention of recordings, transcripts, media, and metadata.	B2, B3, B5
C6	Review and reuse	Sampling of stored recordings and transcripts by machine learning pipelines or trained human reviewers to improve Meta products.	B3

2.3 Privacy Threat Models

The acquisition and movement channels in Section 2.2 tell us how data enters the pipeline and how far it travels, but they do not yet say *who is harmed* or *where the harm originates*. We therefore organize privacy threats along two dimensions: the *data subject* (user or bystander) and the *source of the threat* (system-as-designed or active adversary). This yields

three threat models, which we label TM-1, TM-2, and TM-3 and summarize in Table 3.

TM-1: Platform-Side Threats to the User. This threat covers harms to the user that arise from the device behaving exactly as Meta describes, with no misuse and no attacker. The risk is structural: the pipeline is always listening, retains inadvertent activations, and routes stored audio to human review. For example, living-room television audio is misrecognized as "Hey Meta," the glasses upload roughly ten seconds of surrounding sound that happens to include a private phone call, and because the activation is not flagged as a false wake, the clip is retained and processed in the cloud. Independent research on commercial smart speakers confirms that such misactivations occur regularly in everyday acoustic environments [11, 29]. The threat traverses C1→C4→C5→C6 and terminates not in storage but in human review.

TM-2: Bystander Exposure. This threat covers privacy risks to nearby non-users when the glasses operate as intended. The User controls the glasses, the app, and the account, but nearby people supply the data. Their faces, voices, documents, screens, or behavior can enter the pipeline without giving them any direct notice nor consent. For example, a user enables Live AI for walking directions through a hospital waiting room; for the duration of the session, the faces, voices, and visible intake forms of other patients are processed by Meta's cloud. This risk does not disappear simply because the device provides a capture LED indicator. Indicators may improve notice, but they do not give bystanders control over collection, transfer, retention, review, or reuse [5, 9, 31]. TM-2 spans the widest set of channels (C1–C6) because bystander data can enter at any sensing mode and then traverses the same downstream pipeline as any other captured data, including review and reuse.

TM-3: External-Adversary Threats. The threat originates outside the user–platform–bystander triangle, where an outside adversary exploits the system by either monitoring its visible network traffic or manipulating its exposed AI models. On the network side, the glasses generate a heavily upload-biased traffic profile that is observable from any shared access point [28], and prior work on encrypted-traffic side channels has shown that such patterns are sufficient to distinguish application states and infer user activity [7]; for Meta glasses specifically, this means an observer can infer *when* and *where* the user is actively capturing through Live AI without breaking encryption. On the model side, Chen et al. demonstrated proof-of-concept prompt-injection attacks on Meta Ray-Ban AI glasses, showing that adversarial stickers in the physical environment can hijack the assistant's responses and cause the model to fail at recognizing situational safety hazards [33]. Adversaries with query access to deployed models may further attempt training-data extraction or membership inference against any model that has ingested glasses-collected data [6, 30]. Model inversion is the most consequential variant: Fredrikson et al. reconstructed recognizable training faces from a facial-recognition model [13], which becomes acute if Meta extends the glasses with the facial-recognition feature cited in the Texas Attorney General's recent investigation [25]. TM-3 is the smallest model today and the one whose scope grows fastest with each new feature.

Table 3: Threat models and the channels through which each is realized.

Model	Data subject	Threat source	C1	C2	C3	C4	C5	C6
TM-1	Wearer	System as designed	✓			✓	✓	✓
TM-2	Bystander	System as designed	✓	✓	✓	✓	✓	
TM-3	Wearer or bystander	External adversary				✓	✓	

2.4 LINDDUN Privacy Threat Analysis

In this section, we classify the privacy threats from Section 2.3 using the LINDDUN framework [8, 32], which organizes privacy harms into seven categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure, Unawareness, and Non-compliance. The three threat models told us *who* is harmed and *where* in the pipeline the harm enters; LINDDUN tells us *what kind* of privacy violation each threat is. We summarize the formal definitions of each category in Appendix B.

Table 4 maps each LINDDUN category to a concrete Meta AI glasses threat, its threat model, and the channels through which it arises; here, we focus on the overall patterns revealed by this mapping. From this table, we observe two main findings. **The bystander threat is the most diverse in kind, not just in volume.** TM-2 triggers five of the seven LINDDUN categories (Linkability, Identifiability, Non-repudiation, Unawareness, and Non-compliance) while TM-1 triggers four and TM-3 only two. This means a single Live AI session produces more *kinds* of privacy violation against the bystander than against the user, due to lack of consent. Identifiability and Non-compliance appear only in the TM-2 row, because identifying a person without their permission and processing biometric data without their agreement are violations that only exist when no agreement was given.

Human review (C6) compounds four separate categories. Of the seven LINDDUN categories, Linkability, Non-repudiation, Disclosure, and Unawareness all reach C6. This is what distinguishes Meta AI glasses from smart speakers, which also use human review but do not feed it visual session data: on the glasses, This means a trained reviewer in Meta Cloud opening a single clip, they are seeing a named person, in a specific place, at a specific time, doing a specific

Table 4: LINDDUN privacy threats identified in Meta AI glasses.

Category	Instance on Meta AI glasses	Threat model	Channels
Linkability	Sessions, transcripts, and metadata can be tied to the same user, place, or co-present bystander across time, supporting behavioral profiling and relationship inference.	TM-1, TM-2	C5, C6
Identifiability	Faces, voices, name badges, visible documents, and screen contents in a captured frame can be resolved to a real-world identity by cloud-side vision models.	TM-2	C3, C5
Non-repudiation	A retained transcript or clip creates durable evidence that someone was present, spoke, or acted — including bystanders who never agreed to the record.	TM-1, TM-2	C5, C6
Detectability	An observer on a shared Wi-Fi can detect that a Live AI session is active from the upload-heavy traffic profile [7, 28]; cloud-side models can detect sensitive context (e.g., a hospital, a place of worship) from low-resolution Autocapture frames.	TM-3, TM-1	C4, C5
Disclosure	Inadvertent "Hey Meta" activations cause private speech to be uploaded and made available to trained human reviewers [11, 29]; prompt-injection attacks can cause the assistant to leak captured content to an adversary [33].	TM-1, TM-3	C1, C3, C5, C6
Unawareness	Bystanders cannot distinguish idle eyewear from an active Live AI session; users do not know which inadvertent activations are stored, for how long, or whether their content reaches human reviewers [5].	TM-2	C1, C3, C5, C6
Non-compliance	The device enters spaces (clinics, schools, workplaces, homes) where capture, biometric processing, and retention are constrained by law; bystanders have no practical opt-out and users have no granular control over which channels are active [25].	TM-2	C3, C5, C6

action. Each category alone is a harm; their convergence at a single channel turns C6 into the highest-stakes point in the pipeline. This view tells us where the privacy risks of Meta AI glasses originate and what kinds of harm they produce, but it does not tell us whether users themselves recognize them.

3 Public Privacy Concerns about Meta AI Glasses

To examine which privacy risks users recognize in practice, we conduct a large-scale analysis of online discussions about the glasses on Reddit. We pursue two objectives: to identify which of the risks documented in Section 2 users articulate, and to identify which risks users overlook, particularly those arising at cloud storage and human review. Section 3.1 describes our data collection, and Section 3.2 presents our findings on users' privacy concerns.

3.1 Data Collection

We designed a two-stage data collection pipeline. In the first stage, we searched Reddit posts that were likely to contain user opinions, privacy concerns, or negative perceptions of the device. Specifically, we constructed search queries by combining two product-related terms, 'Ray-Ban Meta" and 'Meta Glass," with three attitude-oriented keywords: 'privacy," 'dislike," and "cons." For each query, our crawler retrieved the top 20 posts. After removing duplicate posts, the dataset consisted of 72 unique posts and 4,900 user comments. In the second stage, we encoded the comments into semantic vectors using the pretrained sentence embedding model `all-MiniLM-L6-v2` [1], and then applied `KMeans` for semantic clustering. We selected $k = 18$ based on exploratory experiments that balanced thematic coherence and interpretability. After clustering, we manually inspected the 18 clusters and selected those directly related to privacy concerns. This filtering step resulted in seven privacy-related clusters, containing 1,771 comments in total.

3.2 Identified Privacy Concerns

To further characterize users' privacy concerns, we applied an LLM-based open coding approach. We prompted GPT-5.4 in a zero-shot setting to infer fine-grained concern categories directly from privacy-related comments, allowing the analysis to capture diverse concerns without relying on a predefined taxonomy. Our classification results reveal two broad sources of privacy concerns around Meta Glass: product-specific concerns and company-level concerns rooted in distrust toward Meta.

3.2.1 Product-Specific Concerns. Among the 1,771 privacy-related comments, 915 reflected product-specific privacy concerns.

Concerns About Unnoticed Recording. One of the most prominent privacy concerns was unnoticed recording. In total, 318 comments discussed this issue, among which 245 explicitly compared Meta AI Glasses with other recording devices, particularly smartphones. For example, one commenter contrasted glasses-based recording with phone-based recording, stating: "It's much better knowing when I'm being recorded than being recorded secretly in a way that's much less conspicuous than someone trying to sneakily record with their phone." This comment suggests that users' concerns were shaped not only by the presence of a camera but also by the device's form factor and social detectability. This comment illustrates that users perceived glasses as more privacy-invasive because recording can be less visible and less socially legible than holding up a phone to record. Beyond detectability, users worried that unnoticed recording could reduce bystanders' agency. As illustrated by Comment 1, when glasses make recording less noticeable,

bystanders may have fewer opportunities to detect, avoid, or object to being captured.

Human Review and Data Processing. A total of 211 comments expressed concerns about how data captured by Meta Glasses would be processed after collection. Some users were concerned that recorded content could be manually reviewed by humans. For example, one commenter stated, "The article seems to indicate that everything you record can be potentially manually reviewed by a human at any point." Another commenter generalized this concern to AI products more broadly, noting that "anything you input in any AI product stands a chance of being reviewed by humans." Other users questioned whether videos captured by the glasses could be transmitted elsewhere for analysis even when the user does not explicitly invoke Meta AI. As one commenter asked, "But if I were to only take a video using them without engaging the Meta AI, will the video be sent elsewhere for data analysis without me knowing?" A further set of users worried that Meta Glasses could serve as a data collection channel for model development, with one commenter stating, "Meta just wants more data to train their AI." These comments show that users' privacy concerns extend beyond immediate recording visibility. Users also questioned the post-collection lifecycle of the data: who can access it, whether it is analyzed by Meta or third parties, and whether it may be repurposed for AI model training.

Reliability of Recording Indicators. Although Meta AI Glasses include an LED indicator intended to signal when recording is taking place, users still questioned whether this signal is sufficiently visible, trustworthy, and resistant to circumvention. In total, 194 comments discussed this issue. Many comments focused on the LED blockers and whether the indicator could be covered, disabled, or otherwise bypassed. For example, Comment 2 illustrates users' concern that there may be practical ways to disable or obscure the recording light. The underlying concern was therefore not simply whether a recording indicator exists, but whether it can function as a reliable and socially visible cue for bystanders.

Boundary of Acceptable Public Recording. Rather than rejecting public recording categorically, users often framed the issue as a boundary problem: when does recording in public become privacy-invasive? 120 comments recognized that people may already be recorded in public environments. However, they distinguished ordinary public recording from recordings that occur without clear notice, awareness, or social visibility. One commenter captured this tension by stating: "Being recorded in public is probably fine, but without knowing you're being filmed, it feels really fxxx up."

Children, Family, and Domestic Contexts. 72 comments discussed privacy concerns involving children, family members, partners, or domestic settings. These comments show that users were concerned not only about being recorded but also about how Meta AI Glasses could enter more intimate social contexts, such as homes, family gatherings, or interactions around children. In these settings, recording may capture people who are unable to provide meaningful consent or may unintentionally capture sensitive information and private moments. For example, Comment 3illustrates that users viewed Meta AI Glasses not only as a personal recording device, but also as a technology that can impose privacy risks on non-consenting bystanders, especially children and family members.

3.2.2 Company-Level Concerns. Beyond product-specific privacy concerns, we found that a substantial portion of users' privacy concerns (407 comments) were rooted in distrust toward Meta.

One source of distrust concerned Meta's business model. Some users framed the glasses as another mechanism for data collection and monetization, rather than merely a consumer device. One commenter wrote, "Oh don't get me wrong, fxxx Meta and the way they sell our info for money. I will never defend them. These glasses are just another tool for them to collect more data from us. I don't trust they're not taking this info and selling it." Here, the concern was not only that the glasses can record the surrounding environment, but that the resulting data may become part of Meta's broader commercial ecosystem.

The other source of distrust came from users' memory of Meta's prior privacy controversies. Comment 4 explicitly referenced *In re Facebook Biometric Information Privacy Litigation*, a lawsuit alleging that Facebook's facial recognition feature collected and stored biometric information from users' photos without obtaining the informed written consent required under the Illinois Biometric Information Privacy Act (BIPA) [2]. The case resulted in a $650 million settlement [3]. This example illustrates that some users' concerns were shaped not only by what the glasses currently claim to do, but also by expectations formed from prior privacy violations.

4 Conclusion

In this paper, we showed that the privacy risks of Meta AI glasses concentrate after capture, in cloud retention, model training, and human review, while public discussion stops at the lens and rarely follows the data any further. Our LINDDUN mapping shows the same pattern in the threat space, since Identifiability and Noncompliance arise mainly where bystanders are never given the opportunity to consent, and these threats are realized downstream, past the point where public attention stops. Appendix D discuss what this means for how privacy is disclosed, signaled, and designed.

Acknowledgment

This work is partially supported by National Science Foundation (NSF) under award numbers 2623260, 2523814, 2448205, and 2239605.

References

[1] [n. d.]. all-MiniLM-L6-v2. https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2.

[2] [n. d.]. Biometric Information Privacy Act. https://www.ilga.gov/Legislation/ILCS/Articles?ActID=3004&ChapterID=57.

[3] [n. d.]. In re Facebook Biometric Information Privacy Litigation. https://www.govinfo.gov/app/details/USCOURTS-cand-3_15-cv-03747/USCOURTS-cand-3_15-cv-03747-16.

[4] Apple. 2026. App Privacy Details on the App Store. https://developer.apple.com/app-store/app-privacy-details/. Accessed: 2026-06-23.

[5] Divyanshu Bhardwaj, Alexander Ponticello, Shreya Tomar, Adrian Dabrowski, and Katharina Krombholz. 2024. In Focus, Out of Privacy: The Wearer's Perspective on the Privacy Dilemma of Camera Glasses. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI)*. 1–18. doi:10.1145/3613904.3642242 Two-week diary study of Ray-Ban Stories wearers (N=15).

[6] Nicholas Carlini, Florian Tramèr, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom Brown, Dawn Song, Úlfar Erlingsson, Alina Oprea, and Colin Raffel. 2021. Extracting Training Data from Large Language Models. In *30th USENIX Security Symposium (USENIX Security)*. 2633–2650.

[7] Mauro Conti, Qian Qian Li, Alberto Maragno, and Riccardo Spolaor. 2018. The Dark Side(-Channel) of Mobile Devices: A Survey on Network Traffic Analysis. *IEEE Communications Surveys & Tutorials* 20, 4 (2018), 2658–2713. doi:10.1109/COMST.2018.2843533

[8] Mina Deng, Kim Wuyts, Riccardo Scandariato, Bart Preneel, and Wouter Joosen. 2011. A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. *Requirements Engineering* 16, 1 (2011), 3–32. doi:10.1007/s00766-010-0115-7

[9] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In situ with bystanders of augmented reality glasses: perspectives on recording and privacy-mediating technologies. In *CHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2377–2386. doi:10.1145/2556288.2557352

[10] Cailbhe Doherty, Maximus Baldwin, Rory Lambe, Marco Altini, and Brian Caulfield. 2025. Privacy in consumer wearable technologies: a living systematic analysis of data policies across leading manufacturers. *npj Digital Medicine* 8 (2025). doi:10.1038/s41746-025-01757-1

[11] Daniel J. Dubois, Roman Kolcun, Anna Maria Mandalari, Muhammad Talha Paracha, David Choffnes, and Hamed Haddadi. 2020. When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers. *Proceedings on Privacy Enhancing Technologies* 2020, 4 (2020), 255–276. doi:10.2478/popets-2020-0072

[12] EssilorLuxottica. 2025. *Q4/Full Year 2024 Results*. Technical Report. https://www.essilorluxottica.com/en/newsroom/press-releases/q4-full-year-2024-results/ Ray-Ban Meta reached 2 million units sold since launch.

[13] Matt Fredrikson, Somesh Jha, and Thomas Ristenpart. 2015. Model Inversion Attacks that Exploit Confidence Information and Basic Countermeasures. In *Proc. ACM SIGSAC Conf. on Computer and Communications Security (CCS)*. 1322–1333. doi:10.1145/2810103.2813677

[14] Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie Cranor. 2023. Speculative Privacy Concerns about AR Glasses Data Collection. *Proceedings on Privacy Enhancing Technologies* 2023, 4 (2023), 416–435. doi:10.56553/popets-2023-0117

[15] Google Play. 2026. Provide Information for Google Play's Data Safety Section. https://support.google.com/googleplay/android-developer/answer/10787469. Accessed: 2026-06-23.

[16] Muhammad Zahid Iqbal and Abraham G. Campbell. 2023. Adopting smart glasses responsibly: potential benefits, ethical, and privacy concerns with Ray-Ban stories. *AI and Ethics* 3 (2023), 325–327. doi:10.1007/s43681-022-00155-7

[17] Marion Koelle, Swamy Ananthanarayan, Simon Czupalla, Wilko Heuten, and Susanne Boll. 2018. Your smart glasses' camera bothers me!: exploring opt-in and opt-out gestures for privacy mediation. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. Association for Computing Machinery, New York, NY, USA, 473–481. doi:10.1145/3240167.3240174

[18] Xiaoan Liu, DaeHo Lee, Eric J Gonzalez, Mar Gonzalez-Franco, and Ryo Suzuki. 2026. VisionClaw: Always-On AI Agents through Smart Glasses. *arXiv preprint arXiv:2604.03486* (2026).

[19] Meta. [n. d.]. Capture photos and videos with AI glasses. https://www.meta.com/help/ai-glasses/272319252352130/. Accessed: 2026-06-24.

[20] Meta. 2024. Ray-Ban Meta Glasses Are Getting New AI Features and Partner Integrations. https://about.fb.com/news/2024/09/ray-ban-meta-glasses-new-ai-features-and-partner-integrations/. Accessed: 2026-06-18.

[21] Meta. 2026. AI Glasses Voice Privacy Notice. Meta Store (Legal). https://www.meta.com/legal/ai-glasses/voice-controls-privacy-notice/ Accessed 2026.

[22] Meta. 2026. How to use Autocapture on AI glasses. Meta AI Glasses Help. https://www.meta.com/help/ai-glasses/1824727564872648/ Accessed 2026.

[23] Meta. 2026. How to use Live AI on AI glasses. Meta AI Glasses Help. https://www.meta.com/help/ai-glasses/894093646030348/ Accessed 2026.

[24] Meta. 2026. Meta AI: Assistant and Glasses. https://apps.apple.com/us/app/meta-ai/id1558240027. Accessed: 2026-06-23.

[25] Office of the Attorney General of Texas. 2026. Attorney General Ken Paxton Launches Investigation Into Meta Glasses to Protect Texans' Privacy From Unlawful Monitoring and Collection of Facial Data. Press release. https://www.texasattorneygeneral.gov/news/releases/attorney-general-ken-paxton-launches-investigation-meta-glasses-protect-texans-privacy-unlawful

[26] Ray-Ban. 2026. Discover Ray-Ban Meta AI Glasses: Specs and Features. https://www.ray-ban.com/usa/l/discover-ray-ban-meta-ai-glasses. Accessed: 2026-06.

[27] Rokid. 2026. Real-Time Translation Glasses: AI-Powered Smart Glasses. https://global.rokid.com/pages/rokid-glasses. Accessed: 2026-06-18.

[28] Sunil Sankaran. 2024. Don't Throw Shade at My Network with Ray-Ban Meta Smart Glasses. VMware VeloCloud Blog. https://blogs.vmware.com/sase/2024/12/05/dont-throw-shade-at-my-network-with-ray-ban-meta-smart-glasses/ Industry traffic measurement report.

[29] Lea Schönherr, Maximilian Golla, Thorsten Eisenhofer, Jan Wiele, Dorothea Kolossa, and Thorsten Holz. 2020. Unacceptable, Where Is My Privacy? Exploring Accidental Triggers of Smart Speakers. *arXiv preprint arXiv:2008.00508* (2020). https://arxiv.org/abs/2008.00508 Extended version in Computer Speech & Language, Elsevier, 2022.

[30] Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. 2017. Membership Inference Attacks Against Machine Learning Models. In *IEEE Symposium on Security and Privacy (S&P)*. 3–18. doi:10.1109/SP.2017.41

[31] Xueyang Wang, Kewen Peng, Xin Yi, and Hewu Li. 2026. Mind the Gap: Mapping Wearer–Bystander Privacy Tensions and Context-Adaptive Pathways for Camera Glasses. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for

Computing Machinery, New York, NY, USA, 28 pages. doi:10.1145/3772318.3791848

[32] Kim Wuyts, Riccardo Scandariato, and Wouter Joosen. 2014. Empirical evaluation of a privacy-focused threat modeling methodology. *Journal of Systems and Software* 96 (2014), 122–138. doi:10.1016/j.jss.2014.05.075

[33] Yicheng Zhang, Zijian Huang, Sophie Chen, Erfan Shayegani, Jiasi Chen, and Nael Abu-Ghazaleh. 2025. Evil Vizier: Vulnerabilities of LLM-Integrated XR Systems. *arXiv preprint arXiv:2509.15213* (2025).

Appendix

A Related Works

Prior work on the privacy risks of AI-enabled glasses has examined specific threat vectors without tracing the complete data pipeline. Studies targeting Meta's glasses and their predecessors have identified ethical and social harms that arise from unconsented recording in public spaces and data collection within Meta's ecosystem [14, 16]. Research on AR glasses data collection has further shown that users anticipate risks of data aggregation, secondary use, and cloud transfer, yet these studies rely on speculative user interviews rather than auditing what manufacturers formally document [14]. Systematic audits of consumer wearable privacy policies have demonstrated widespread inconsistency in how manufacturers disclose retention practices and third-party data sharing [10], but none of these audits trace the full data lifecycle of an AI glasses product from the moment of capture through cloud processing and human review. Research on adjacent always-listening devices has independently confirmed that smart speakers misactivate regularly in everyday environments and that the resulting audio reaches cloud storage and human reviewers without user awareness [11, 29], a finding our paper extends to Meta AI glasses by grounding the analysis in Meta's own documentation. Our study advances beyond all of these by applying the LINDDUN privacy threat framework [8, 32] to the Meta AI glasses data pipeline, which allows us to name precisely what category of violation each pipeline stage produces rather than characterizing risk at the level of the device as a whole.

Research on how users and bystanders perceive privacy risks from wearable cameras has concentrated overwhelmingly on the act of capture rather than on the data practices that follow it. Diary studies with Ray-Ban camera glasses wearers have established that users consider existing capture indicators ineffective and feel personally responsible for protecting bystanders in the absence of adequate technical safeguards [5]. Bystander-centered studies have demonstrated that people near smart glasses cannot distinguish active recording from idle use, and that this perceptual ambiguity generates persistent anxiety about unconsented capture in public and sensitive spaces [9, 31]. Studies examining the social acceptability of data glasses confirm that public concern tracks the visible act of recording rather than the less visible downstream data practices that follow it [17]. Together, these studies establish that user concern concentrates on the camera lens and diminishes further down the data pipeline. However, no prior work directly analyzes publicly expressed privacy concerns around Meta AI glasses at scale.

B LINDDUN Threat Category Definitions and Per-Category Analysis

Linkability. The risk that distinct data points are connected to the same person, place, or relationship across sessions. A user who repeatedly uses Live AI or Autocapture in fixed settings (a workplace, a clinic, a campus) accumulates context in which the same bystander recurs. Timestamps, location metadata, and visual context can then link that bystander to the user and reconstruct an offline relationship the bystander never disclosed. The same linkage also enables longitudinal profiling of the user across reviewed sessions at C6. (TM-1, TM-2; C5, C6.)

Identifiability. The risk that a real-world identity is inferred from captured data. A bystander's face, a visible name badge, a laptop screen, or a medical form entering the user's field of view becomes input to a visual model capable of OCR and scene understanding (B5). Identification need not be the user's goal; it is a property of the processing pipeline. (TM-2; C3, C5.)

Non-repudiation. The risk that the system creates durable evidence that someone was present, spoke, or acted. A cloud transcript or clip can fix on record that a person attended a protest, sat in a medical waiting room, or participated in a private conversation, evidence the subject did not choose to create and cannot retract. Once such evidence enters human review at C6, it is no longer recoverable through deletion. (TM-1, TM-2; C5, C6.)

Detectability. The risk that a sensitive fact becomes observable to an actor. Three distinct vectors arise: (a) a network observer distinguishes an active Live AI session from idle eyewear by traffic volume and timing (TM-3, C4); (b) the cloud model infers from audio and low-resolution frames that the user is inside a hospital or a place of worship (TM-1, C5); (c) OCR detects that a document in view is a legal or medical form (C5). In each case, a previously hidden fact is surfaced to a non-subject. (TM-1, TM-3; C4, C5.)

Disclosure. The risk that sensitive data is exposed to an unintended party. Meta states that inadvertent activations are stored (B2) and that stored audio may be heard by trained reviewers (B3); prior smart speaker research independently shows accidental triggers uploading up to a minute of unintended speech. Ambient private speech can thus travel from C1 all the way to a human reviewer at C6. Disclosure also arises through external manipulation: a visual prompt injection during a Live AI session (C3) can cause the assistant to

reveal captured content to an attacker. (TM-1, TM-3; C1, C3, C5, C6.)

Unawareness. The risk that the data subject lacks notice, comprehension, or control. A capture LED communicates a single bit, recording or not, but the device occupies at least six privacy-relevant states: idle, wake-word listening, explicit capture, Live AI, Autocapture, and cloud upload. A bystander can distinguish none of these states from any other, has no interface to inspect what was captured about them, and has no path to request deletion once data enters cloud storage at C5 or human review at C6. (TM-2; C1, C3, C5, C6.)

Non-compliance. The risk that the system conflicts with consent, biometric, retention, or sensitive space obligations. The device's mobility carries it into clinics, schools, workplaces, and homes where recording and biometric processing norms impose legal duties, yet bystanders have no practical mechanism to consent or opt out. Compliance risk extends into review and reuse at C6, where biometric material from non-consenting subjects may be sampled for model improvement. We frame specific statutory exposure as a research question rather than a finding, since it requires legal adjudication (TM-2; C3, C5, C6.).

Table 5: LINDDUN privacy threat categories.

Category	Definition
Linkability	Distinct data items can be tied to the same person, device, session, or context, enabling profiling across time even when each item alone is anonymous.
Identifiability	A data subject's real-world identity can be inferred from the data, directly or through combination with other available information.
Non-repudiation	The system creates a durable record that a data subject was present, said something, or performed an action, such that the subject cannot plausibly deny it.
Detectability	The existence of a sensitive data item, event, or system state becomes observable to an actor (a platform, model, reviewer, or external observer), even when its content is not disclosed.
Disclosure	Sensitive data is exposed to an unintended or unauthorized party through storage, transmission, processing, sharing, or model output.
Unawareness	The data subject lacks notice, comprehension, or control over what is being collected, processed, retained, reviewed, or reused about them.
Non-compliance	The system's data practices conflict with applicable consent requirements, retention limits, biometric protections, or sensitive-space obligations under law or policy.

C Representative Comments

Section 3.2 refers to three Reddit comments that we quote in full here, duo to page limit. We selected them from the 1,771 privacy-related comments as the main concern categories: unnoticed recording, the reliability of the capture indicator, and capture in domestic settings involving children. Each comment is quoted verbatim except that profanity is masked, and we report don't usernames, post titles, or links, since our aim is only to characterize how privacy concerns are expressed.

Comment 1

"They wouldn't have known if this mom didn't point it out. If you're recording the game on your phone, it's pretty obvious. It also gives you, as the parent, the opportunity to stop the recording while it's happening if it makes you uncomfortable. You can't just go up to everyone wearing glasses and say hey don't watch my kid."

Comment 2

"I work in an underwear/sleepwear store, and I've kicked men out for having them on. I don't want to hear their 'but they're not recording right now' excuse either, I know there are plenty of ways to disable the light from a piece of dark tape to actually overriding it. There are so many places where being able to film discreetly is a HORRIBLE idea."

Comment 3

"forWhat you are missing is that it's not all about you. You made a choice, but by your own admission, you spend significant time around children who can not give consent. At best, from your responses, you have a concerning disregard about potentially putting children's privacy at risk so you can have a 'cool' but unnecessary wearable, at worst, it could be considered negligent."

D Discussion

The gap between where the risks of Meta AI glasses concentrate and where public concern settles has direct consequences for how privacy is disclosed, signaled, and designed. Disclosure has to reach users in real time and account for the full data lifecycle, rather than resting in privacy policies that users rarely read. Signaling has to move past the single recording LED toward multimodal notification, and, because notification will sometimes fail, toward automated on-device safeguards such as blurring faces and screens before data leaves the glasses. Meta must stop treating the wearer as the only customer, and instead treat bystanders, household members, and people in sensitive settings such as clinics and schools as equal stakeholders.