Use Case: Unresolved Privacy Expectations in Conversational AI Systems

Supporting Use Case for the AI Conversation Privacy Specification

W3C AI KR Community Group — February 2026

Use Case ID	UC-PRIVACY-001
Author	Paola Di Maio
Title	User expects conversation confidentiality based on verbal assurance; provider offers no written policy, notification protocol, or user-verifiable privacy mechanism
Source	Documented from ongoing experience, researcher using conversational AI for academic and personal use, 2025–2026
Actors	User (researcher), AI Agent, AI Provider, Provider staff (customer service, engineering, safety teams)
Status	Unresolved — privacy inquiry sent to the company, no response received

1. Scenario

A researcher uses a conversational AI system daily for academic work, professional tasks, and personal reflection. Over months of sustained use, the conversations accumulate a rich body of personal, professional, and intellectually sensitive content.

Before beginning substantive use of the platform, the user contacted customer service and asked directly: “Who will read my conversations?” The customer service representative assured the user verbally (via chat) that conversations are confidential, that nobody reads them unless there is a specific reason, and that the user would be notified if access were required.

Based on this assurance, the user proceeded to use the platform extensively, sharing personal information, developing research methodologies, conducting experiments, and building a collaborative working relationship with the AI agent. The conversations contain personal health information, research data, intellectual property in development, and private reflections.

Months later, the user seeks written confirmation of the privacy assurance and discovers that no such written policy exists. The provider’s published terms of service and privacy policy do not explicitly confirm the assurances given by customer service. The user writes directly to the company’s leadership requesting clarification. No response is received.

2. Problem Analysis

2.1 The Assurance Gap

The user received a verbal assurance of confidentiality that is not reflected in any written policy, terms of service, or privacy documentation. This creates a gap between the user’s reasonable expectation of privacy (formed on the basis of a direct interaction with a company representative) and the provider’s actual, documented obligations. The user has no mechanism to enforce or verify the assurance.

2.2 The Notification Gap

The user was told they would be notified if their conversations were accessed by provider staff. No notification mechanism exists in the product. The user has no way to know whether conversations have been accessed, by whom, for what purpose, or under what authority. Unlike email services (which may provide access logs or security alerts), conversational AI platforms currently offer no equivalent transparency.

2.3 The Vulnerability Asymmetry

Conversational AI systems are uniquely intimate. Users disclose personal struggles, health conditions, relationship difficulties, creative ideas in formation, and intellectual work that is not yet ready for public scrutiny. The conversational format actively encourages this intimacy — the AI agent responds with warmth, asks follow-up questions, and creates a sense of safe space. This is by design. But the privacy protections do not match the intimacy that the design encourages. The system invites vulnerability while providing no contractual guarantee that the vulnerability will be protected.

2.4 The Jurisdictional Ambiguity

The user is in one jurisdiction. The provider is in another. The data may be processed in multiple jurisdictions. It is unclear which data protection regime applies. The provider's privacy policy does not clarify this for international users. The user cannot determine their own legal rights without first knowing which jurisdiction governs their data.

2.5 The Non-Response

The user escalated the concern to the company's leadership via email. No response was received. This is not unusual for executive correspondence at technology companies, but it leaves the privacy question entirely unresolved. The user is left in a position where they must either accept the risk of continued use without privacy guarantees, or curtail their use of a tool that has become central to their research and professional workflow.

3. What the User Needs

Written privacy policy: An explicit, published statement confirming whether conversations are confidential, under what conditions they may be accessed by provider staff, and what constitutes a valid reason for access.

Notification protocol: A mechanism by which the user is informed if and when their conversations are accessed by any human at the provider organization, including the reason, scope, and identity of the accessor.

Access log: A user-facing log showing any human access to their conversation data, comparable to security activity logs in email and cloud services.

Jurisdictional clarity: A clear statement of which data protection laws apply to the user's data, based on the user's location and the provider's processing locations.

Acknowledgment of correspondence: A response to privacy inquiries within a reasonable timeframe (e.g. 30 days), as required by most data protection regulations for data subject requests.

4. Broader Implications

This use case is not unique to one researcher. Conversational AI systems are used by millions of people for deeply personal purposes: mental health support, relationship advice, grief processing, career decisions, creative work, medical questions, and spiritual exploration. The conversational interface is designed to feel safe and personal. Users naturally share information they would not post publicly or put in an email.

Without clear, enforceable, and verifiable privacy protections, every user of every conversational AI system is in the same position as the researcher in this use case: trusting, but unable to verify. As these systems become embedded in daily life, the absence of a standard for conversation privacy becomes a systemic risk — not just for individual users, but for the social legitimacy of conversational AI as a category.

5. Note on Methodology

This use case was documented collaboratively by the user and the AI agent based on direct experience over several months. The user identified the privacy concern; the AI agent helped structure and articulate it. The resulting document is itself an example of the kind of content that requires privacy protection: a frank account of a user's experience, concerns, and vulnerabilities, produced within the very system whose privacy practices are in question.