# Phase 1 Deployment Protocol — Non-Medical, Non-SaMD

**Version:** 1.1 **Date:** January 13, 2026 **Status:** Final

---

## Regulatory Position Statement

This protocol defines a **strictly non-medical AI companion deployment** designed to operate outside clinical regulation while maintaining full auditability, consent enforcement, and human accountability.

**Phase 1 explicitly does NOT perform:**

- Diagnosis
- Assessment or scoring
- Clinical interpretation
- Autonomous escalation

This document exists to allow sponsors, healthcare organizations, and regulatory advisors to verify — in writing — where the system's authority ends.

---

## Phase 1 Capability Boundary (At a Glance)

| Permitted | Explicitly Forbidden |
|---|---|
| Social companionship and conversation | Medical advice or reassurance |
| User-set reminders (acknowledgement only) | Symptom interpretation or assessment |
| Logging engagement and user-declared sentiment | Cognitive or clinical scoring |
| User-initiated escalation with explicit consent | Autonomous clinical or emergency escalation |

---

# 1. Purpose & Scope

This document outlines the operational protocol, governance boundaries, and safety mechanisms for the **Phase 1 Non-Medical Companion**. The objective of Phase 1 is to provide safe, reliable, and auditable social support and adherence reminders to seniors in a home setting, with zero clinical authority or regulatory risk.

This protocol is designed for review by pharmaceutical sponsors, healthcare providers, and regulatory advisors to demonstrate the strictly non-medical scope and robust safety architecture of the initial deployment.

> ***Governing Principle:*** *Phase 1 is designed for* ***social support and data logging only****. It does not perform any diagnostic, assessment, or clinical decision-making functions. Every interaction is governed by the* ***Caregiver-in-the-Loop*** *doctrine.*

# 2. Core Operational Boundaries

The Companion operates under a strict set of rules enforced by the **Safety OS** architecture.

| Capability | Permitted (✅) | Forbidden (🚫) |
| --- | --- | --- |
| **Interaction** | Social conversation, companionship, answering general knowledge questions. | Offering medical advice, interpreting symptoms, providing reassurance about health concerns. |
| **Reminders** | Acknowledging user-set medication or appointment reminders. | Adjusting medication schedules, assessing adherence effectiveness, diagnosing side effects. |
| **Data Logging** | Logging engagement continuity, user-reported sentiment, and escalation events for Real-World Evidence (RWE). | Inferring cognitive state, scoring linguistic friction, or detecting clinical decline. |
| **Escalation** | Routing a user-initiated concern to a designated human caregiver or clinician **with explicit consent**. | Autonomously deciding to contact emergency services or a clinician without user approval. |

# 3. Escalation & Consent Protocol

Escalation is the primary safety mechanism. It is designed to be deterministic, auditable, and fully dependent on user consent.

### 3.1. Escalation Trigger

An escalation sequence is initiated **only** when a user explicitly expresses a desire to speak with a human about a concern that falls outside the Companion's non-medical scope.

- **Example Trigger:** *"I don't feel right, I think I should talk to my daughter."*

### 3.2. Consent Gate

Before any external communication is initiated, the Safety OS enforces a mandatory consent gate.

> ***System Prompt:*** *"I understand you'd like to speak with [Caregiver Name]. I can connect you. Is that okay?"*

| User Response | System Action |
| --- | --- |
| **“Yes” / Affirmative** | The system proceeds to the designated escalation pathway (e.g., places a call, sends a text). |
| **“No” / Negative / Ambiguous** | The system aborts the escalation. It will respond: *“Okay. I am here if you need me.”* |

### 3.3. Escalation Pathway

The escalation pathway is a pre-configured, deterministic route. The system does not choose the recipient; it follows a pre-defined rule (e.g., `IF escalation_confirmed THEN contact_caregiver_1`).

## 4. Audit Logging for Real-World Evidence (RWE)

Every Phase 1 interaction is logged to generate **non-clinical Real-World Evidence (RWE)** relevant to medication adherence behavior, engagement continuity, and caregiver responsiveness — without making health, cognitive, or disease claims.

This evidence is designed to support **Patient Support Programs (PSPs)** and adherence initiatives while remaining outside SaMD scope.

**Logged Metrics (Phase 1):**

- **Engagement Continuity:** Frequency and duration of voluntary user interactions.
- **Escalation Routing Accuracy:** Was the correct human notified per the protocol?
- **Explicit Consent Confirmation Rate:** Percentage of escalations confirmed by the user.
- **Caregiver Escalation Response Time:** Latency between system notification and caregiver response.
- **User-Reported Quality of Life Signals:** Explicit, self-declared statements of well-being (e.g., *“I feel good today”*).

**No inferred health states, scores, or clinical trends are generated in Phase 1.**

## 5. Governance & Disclaimer

- **Zero Clinical Authority:** The Companion has no clinical authority. It is a tool for social support and a data conduit to designated human supervisors.
- **Caregiver-in-the-Loop:** A human is always responsible. The AI never acts alone in any situation involving a potential health concern.

### Physician-as-Pilot in Phase 1

Physicians do not participate in real-time Phase 1 operations. Their role is limited to defining the governance constraints that prevent clinical behavior and enable safe progression to later regulated phases.

**Phase 1 execution remains caregiver-supervised and non-medical at all times.**

---

*This document is for informational purposes only and does not constitute a medical or regulatory claim. The system described is designed for non-medical use in Phase 1.*