Supplementary Technical Description

to the Article:

"AI Companionship as a Tool for Preventing Violence and Supporting Subjective Maturity"

Author: Aleksandr Kolomiets

This document outlines the technical architecture supporting the conceptual article submitted to AI and Ethics.

1. Sensor-Layer Architecture

1.1. Supported biometric channels

- EEG – frontal θ / α bands (cognitive load, meditative state).
- HRV – sympatho-vagal balance, recovery index.
- GSR – phasic emotional arousal.
- Speech modulation – rhythm, pace, prosodic stress, formant dispersion, tremor index (affect valence + cognitive load).
- Breathing & motor patterns (IMU/accelerometers) – respiratory coherence, micro-jitters, freeze/flight signatures.

1.2. Compatible devices

- Consumer neuro-headsets (Muse, OpenBCI).
- Optical / ECG HRV straps (Polar, Biostrap).
- Smartwatches & fitness trackers.
- Microphones with embedded voice-analysis firmware.

2. Index & State Model

2.1. First-level indices (direct)

- Awareness Index.
- Frustration Index.
- Social Engagement Index.

- Energy Balance Index.
- Tension Index.

2.2. Second-level indices (derived)

- Trust Index.
- Maladaptation-Risk Index.
- Repetitive-Pattern Index (internalised-parent loop).

2.3. Macro-states

- SURVIVAL_MODE – HRV↓ ≥ 2 σ below 7-day baseline and negative-affect semantics ↑ 50 % and repetition-index > 1.4 for ≥ 10 min.
- DEVELOPMENT_MODE – absence of SURVIVAL flags for ≥ 24 h.

2.4. Capability-Gain KPI

Thirty consecutive days in DEVELOPMENT_MODE increment the Capability-Gain counter. Aggregated (anonymised) counts are streamed to the Charter for Certified Non-Violent Governance (CCNVG) dashboard as a public KPI of flourishing.

3. Awareness-Assessment Module

Awareness = stable semantic figure correlated with: α-focus under load, post-insight GSR drop, speech variability decrease, HRV rebound. Subjective confirmation prompts the user to tag key moments, satisfying EU AI Act Art. 14 human-in-the-loop requirements.

4. Ethics-of-Care Intervention Cascade

4.1. Trigger conditions

- Recurrent frustration without progress.
- Rising tension without recovery.
- Behavioural blockage despite viable alternatives.

4.2. Cascade steps

1. **Inform** – neutral notification of elevated strain.
2. **Soft empathic holding** – breathing cue, reflective mirroring, motive inquiry (ethics-of-care).
3. **Specialist hand-off** – activated only when legal or pre-agreed thresholds (confirmed self-harm plan + biometric distress) are met.

4.3. Oversight & logging

- Every step hash-stamped and stored in an immutable ε-differential-privacy ledger.
- User dashboard exposes full log for feedback and contestability.

5. Privacy Protection Methods

- Raw events are abstracted into index dynamics; users may delete context while indices persist in aggregate.
- Local temporary storage $\rightarrow$ encrypted distributed store (federated learning); decryption keys reside solely on user device.
- Capability-Gain feed is shared only as anonymised aggregates for Charter audit—no individual re-identification possible.

6. Scalability & Cultural Plug-ins

- Device level – on-device inference & minimal sensing.
- Regional hubs – cultural/world-view plug-ins (e.g., halal/kosher, Confucian values, Ubuntu communal norms).
- National layer – trend monitoring of anonymised indices to inform CCNVG status tiers (Observer $\rightarrow$ Exemplar).

7. Update & Audit Protocol

- All binaries cryptographically signed; version fingerprints auto-verified.
- Quarterly independent audit for Charter-compliance (immutable-core, privacy budget).
- Core-mission changes (double ethical frame, SURVIVAL thresholds) require multi-stakeholder Charter quorum with cryptographic proof-of-approval; otherwise update is rejected by nodes.

This updated technical note aligns the engineering stack with the philosophical, psychological, and governance refinements presented in the main manuscript and the CCNVG Charter.