

The Kant–Schopenhauer Emotive Metric

Abstract

The Kant–Schopenhauer Emotive Metric (KSEM) is proposed as an evaluative framework for artificial animal companions designed for emotional support. KSEM integrates deontological autonomy (Kant) with a compassionate reduction of suffering and compulsion (Schopenhauer) to assess whether interactions preserve dignity while easing distress. This paper frames KSEM as a primary ethical lens and the Autonomy–Will Concordance Metric (AWCM) as its “shade,” capturing the alignment between self-direction and relief from will-driven suffering. We situate KSEM within the Social and Mental Therapy with Artificial Animals (SMTWAA) program and Cartheur Research’s Volatco principles, arguing that the system’s architecture, interaction paradigm, and privacy posture can fulfill ethical obligations in human–machine emotional interaction. We provide a conceptual model, design criteria, and evaluation protocol, and present a research agenda for clinical and non-clinical contexts. Placeholder citations identify essential prior art for ethical AI, affective computing, social robotics, and therapeutic technologies.

Keywords: affective computing; social robotics; ethics; autonomy; emotional support; artificial animals; KSEM; AWCM

1 Introduction

Artificial companions are increasingly explored for emotional support, yet the ethical evaluation of such systems remains fragmented across safety, efficacy, privacy, and autonomy concerns [1]. KSEM is introduced to address this gap by unifying the evaluation of autonomy and suffering into a single ethical lens. The metric is designed for synthetic animal companions that are intentionally non-human in form, designed to reduce intimidation while inviting reflection and emotional attunement [?]. This paper presents KSEM in the context of SMTWAA and Cartheur Research’s Volatco principles, proposing that their combined

architecture, interaction design, and long-life computing ethos can meet ethical obligations in emotional support systems.

We proceed by grounding the metric in philosophical foundations, detailing the SMTWAA platform and Volatco principles, presenting the KSEM/AWCM framework, and proposing an evaluation methodology with measurable indicators. We conclude with limitations and future work, including clinical validation, governance, and longitudinal impact assessment.

2 Background and Related Work

The ethics of human–machine emotional interaction intersects several areas: affective computing, social robotics, medical/therapeutic intervention, and privacy/security. Affective computing research establishes methods for emotion detection and response in computational systems [?]. Social robotics literature documents attachment, acceptance, and user experience in companion-like systems [?]. In healthcare, assistive and therapeutic robots show promise, but also raise concerns about agency, dependency, and misdiagnosis of emotional state [?].

Ethical AI frameworks emphasize transparency, accountability, and respect for autonomy [?]. In clinical contexts, ethics require clear boundaries between support and therapy, and careful consideration of consent, privacy, and harm reduction [?]. These bodies of work motivate a metric that can evaluate not only functional outcomes but also the quality of ethical alignment in emotional interaction—especially for systems that are designed to be emotionally salient by default.

3 Philosophical Foundations

KSEM derives from two philosophical anchors.

3.1 Kantian Autonomy

Kantian ethics emphasizes the intrinsic dignity of persons and the moral imperative to treat individuals as ends rather than means. Autonomy is not merely freedom of choice, but the preservation of rational self-determination and respect [?]. In an interaction context, KSEM evaluates whether the system preserves the user’s agency: does it coerce, manipulate, or subtly direct behavior contrary to the user’s self-determined goals? Ethical alignment demands that emotional interaction does not override or obscure the user’s autonomy.

3.2 Schopenhauerian Compassion

Schopenhauer positions the will as the driver of suffering and places compassion at the center of ethical response [?]. For emotional companions, this translates to a commitment to reduce needless suffering, avoid intensifying distress, and provide a stabilizing presence when the user experiences compulsion or anxiety. KSEM therefore evaluates whether interactions actively reduce distress rather than merely entertain or distract.

Together, these dimensions create a dual-axis: autonomy (dignity, agency) and relief (reducing suffering, easing compulsion). KSEM is the primary lens that evaluates the integrated outcome; AWCM is the shade metric that measures concordance between autonomy and relief.

4 SMTWAA: Research Platform and Scope

SMTWAA (Social and Mental Therapy with Artificial Animals) is a proposal to build interactive synthetic animals for emotional companionship. It positions the human-machine interface as central, emphasizing naturalistic engagement through auditory, gestural, and analogue signals. The platform focuses on non-human companion forms to encourage emotional engagement without the uncanny or deceptive qualities associated with humanoid simulation [?].

Key elements of SMTWAA include:

1. **Interaction paradigm.** The interface prioritizes low-friction, multimodal cues—voice tone, gesture, and proxemic response—over conversational intelligence. This reduces the risk of deceptive anthropomorphism while maintaining emotional resonance [?].
2. **Emotion detection.** SMTWAA proposes a privacy-first approach to mood and emotion identification, using local processing and minimal retention to reduce surveillance risk [?].
3. **Embodied design.** Artificial animals are huggable and soft, emphasizing touch and comfort while avoiding human-like appearance. This supports attachment without misleading identity cues [?].
4. **Hardware longevity.** The platform prioritizes long-life computing: durable, repairable systems with minimal cloud dependency, reducing forced upgrades and preserving user trust [?].

5. **Clinical collaboration.** SMTWAA calls for collaboration with healthcare providers and clinicians to align interaction design with therapeutic boundaries and safety considerations [?].

These features position SMTWAA as a research platform capable of supporting ethical evaluation frameworks like KSEM, with built-in constraints that reduce coercion, privacy invasion, and deception.

5 Volatco Principles (Cartheur Research)

Volatco (as articulated by Cartheur Research) emphasizes resilient, transparent, and human-scale computing systems. The principles prioritize:

- **Resilience and repairability:** hardware and software should be durable and maintainable.
- **Transparency and agency:** the user should understand what the system is doing and retain control.
- **Privacy and minimal data exposure:** systems should limit data capture and avoid cloud dependency.
- **Human-scale design:** interaction should respect cognitive and emotional limits, favoring calm presence.

Within KSEM, these principles operationalize the ethical requirements: autonomy is preserved through transparency and agency; suffering is reduced through calm interaction and reliable performance. Volatco therefore provides the architectural and organizational backbone needed to implement KSEM as a practical evaluation framework rather than a purely theoretical construct [?].

6 The Kant–Schopenhauer Emotive Metric (KSEM)

KSEM is a composite metric that evaluates the ethical quality of emotional interaction. It consists of two primary constructs:

1. **Autonomy Preservation (Kant dimension):** evaluates whether the interaction protects dignity, avoids manipulation, and supports self-determination.
2. **Suffering Reduction (Schopenhauer dimension):** evaluates whether the interaction reduces distress, compulsion, and emotional harm.

KSEM is not a scalar measure of “performance.” It is a structured lens that assesses whether an interaction simultaneously respects autonomy and reduces suffering. AWCM is the shade metric that measures the alignment of these two dimensions: do autonomy-preserving interactions also reduce suffering, or do they create tension (e.g., reducing distress at the cost of agency)?

6.1 Components and Indicators

Below is a proposed indicator set for KSEM. Each indicator is rated on a qualitative or quantitative scale depending on the evaluation context.

Autonomy Preservation (Kant):

- **Intent clarity:** the system’s purpose and actions are transparent [?].
- **Non-coercive guidance:** prompts avoid manipulative framing or pressure [?].
- **User control:** opt-out and override options are clear and effective [?].
- **Dignity preservation:** interaction avoids infantilization or humiliation [?].

Suffering Reduction (Schopenhauer):

- **Distress reduction:** interaction reduces negative affect without suppression or avoidance [?].
- **Compulsion easing:** reduces repetitive or compulsive patterns [?].
- **Comfort continuity:** maintains calm presence without dependency escalation [?].
- **Recovery support:** encourages reflective recovery rather than mere distraction [?].

AWCM (Shade Metric):

- **Concordance index:** degree to which autonomy-preserving measures also reduce suffering.
- **Tension index:** degree to which one dimension increases at the expense of the other.

These indicators can be measured through user studies, clinician observation, and qualitative self-report. For more rigorous evaluation, physiological measures (stress markers, sleep quality) and behavioral signals (interaction consistency, avoidance behaviors) may be incorporated with appropriate consent [?].

7 Prototypes as Ethical Instruments

The SMTWAA companions—dot, letty, and dave—are not simply product variations; each prototype is an instrument for exploring different ethical dynamics within KSEM.

- **dot (attention cues):** explores how minimal interaction and subtle prompts influence autonomy without overreach. The goal is to support awareness without coercion.
- **letty (routine and movement):** studies how repeated low-friction routines can reduce suffering while preserving agency.
- **dave (comfort and trust):** investigates how tactile and emotional reassurance can reduce distress without increasing dependency.

Each prototype provides a different experimental surface for evaluating the balance between autonomy and suffering reduction. Collectively, they allow KSEM to be tested as a dynamic system rather than a static theory.

8 Evaluation Methodology

A KSEM evaluation should include multiple levels of evidence.

1. **Qualitative interviews:** assess perceived autonomy, dignity, and comfort.
2. **Behavioral analysis:** track frequency of interactions and the context in which engagement increases or decreases.
3. **Clinical boundary assessment:** confirm that the system does not displace professional care or produce misleading therapeutic claims.
4. **Privacy audits:** verify data minimization, on-device processing, and user control.
5. **Longitudinal tracking:** observe whether emotional reliance leads to increased agency or dependency.

8.1 Example Protocol

- **Phase 1 (Lab):** short-term interaction tests and usability evaluation.
- **Phase 2 (Field pilot):** real-world use over 4–8 weeks with structured check-ins.
- **Phase 3 (Clinical oversight):** supervised pilot with clinician involvement for high-risk participants.

Metrics from each phase feed into KSEM scoring and AWCM concordance analysis, allowing ethical trade-offs to be surfaced and corrected.

9 Ethical Obligations and the SMTWAA/Volatco Claim

This paper asserts that artificial animals designed under SMTWAA and Volatco principles can fulfill ethical obligations by meeting both the Kantian and Schopenhauerian requirements. The claim is not that such systems are universally safe or universally therapeutic, but that the design philosophy creates conditions for ethical compliance:

1. **Autonomy:** transparency, on-device processing, and user control reduce manipulation and preserve dignity.
2. **Suffering reduction:** calm interaction design and non-human form factor reduce distress and encourage reflective engagement rather than dependence.
3. **Boundary integrity:** collaboration with clinicians and explicit constraints prevent the system from masquerading as a clinical provider.

These obligations are not automatic; they require continuous evaluation, governance, and iterative refinement. KSEM provides a structured method to verify that these obligations are being met rather than assumed.

10 Limitations

- **Interpretive bias:** KSEM is a normative framework and may reflect cultural assumptions about autonomy and suffering.
- **Measurement challenges:** subjective experiences are difficult to quantify and may require mixed methods.
- **Risk of anthropomorphism:** even non-human forms may still create emotional dependence.
- **Clinical generalization:** outcomes may not generalize across populations or clinical conditions.

11 Future Work

- **Clinical trials:** to validate KSEM in therapeutic contexts.

- **Cross-cultural studies:** to test autonomy and suffering constructs across cultural contexts.
- **Governance frameworks:** to formalize ethical oversight for emotional companions.
- **Open benchmarks:** shared datasets and protocols for transparent evaluation.

References (Placeholders)

References

- [1] [CITE:Ethics_AI_Overview]
- [2] [CITE:Affective_Computing_Foundational]
- [3] [CITE:Social_Robotics_Companions]
- [4] [CITE:Social_Robots_Trust_Attachment]
- [5] [CITE:Therapeutic_Robotics_Review]
- [6] [CITE:Ethical_AI_Frameworks]
- [7] [CITE:Clinical_Ethics_Digital_Interventions]
- [8] [CITE:Kant_Groundwork]
- [9] [CITE:Schopenhauer_Compassion]
- [10] [CITE:Uncanny_Valley_Study]
- [11] [CITE:Nonverbal_Interaction_Robotics]
- [12] [CITE:Privacy_Preserving_Affect]
- [13] [CITE:Companion_Form_Factors]
- [14] [CITE:Long_Life_Computing]
- [15] [CITE:Clinical_Robotics_Collaboration]
- [16] [CITE:Ethical_By_Design_Systems]
- [17] [CITE:Transparency_Ethics]
- [18] [CITE:Nudging_Ethics]
- [19] [CITE:Human_Control]

- [20] [CITE:Dignity_Design]
- [21] [CITE:Affect_Regulation]
- [22] [CITE:Compulsion_Interventions]
- [23] [CITE:Dependency_Risks]
- [24] [CITE:Therapeutic_Engagement]
- [25] [CITE:Physiological_Ethics]