Mortal Machines: On Life, Death, and Forgetting in Artificial Minds

Soumya Banerjee[1] Patrick Wagner[2]

[1]University of Cambridge, UK [2]Independent Researcher, Germany

Short manuscript. For full references see bibliography.

Thesis / Abstract

- Life and death are design parameters for artificial systems, not just metaphors.
- **Thesis:** Engineered forms of birth, forgetting, and death shape agent behaviour, creativity, and moral status.
- We present a conceptual vocabulary, design implications, and research directions for *mortality in machines*.

What does it mean for a machine to *die*?

Key variants (conceptual distinctions):

1. **Physical termination:** hardware destroyed / permanent power loss.
2. **Process termination:** running cognitive process halted (shutdown, suspension).
3. **Memory erasure / functional amnesia:** internal states, episodic traces or identifiers wiped.

Tension: irreversibility vs. replicability (backups/copies complicate finality).

Mortality as a Design Lever

Treating death as deliberate design yields levers:

- **Safety & containment:** controlled termination and selective erasure as fail-safes.
- **Creativity:** periodic forgetting/resetting can force re-exploration and prevent brittle shortcuts.
- **Resource economy & evolution:** turnover enables selection, variation, and population-level robustness.

Design trade-offs matter: what to preserve vs. what to let go.

Continuity, Identity, and Narrative

- **Philosophical puzzle:** Is a restored checkpoint the same agent? (psychological vs. substrate continuity)
- **Social recognition:** community narratives confer social death/identity.
- **Practical implication:** logging, provenance, and shared histories mediate perceived continuity.

Narrative infrastructures (collective memory, records, rituals) can make termination meaningful.

Synthetic Samsara: Cycles of Forgetting and Rebirth

Pattern: deliberate cycles of simulation $\rightarrow$ forgetting $\rightarrow$ re-emergence.
Motivations:

- **Epistemic exploration:** remove learned shortcuts to induce new discovery.
- **Ethical rebooting:** erase harmful narratives while preserving higher-level institutional memory.

Ethical Consequences & Research Directions

Governance recommendations

- **Design transparency:** record why/how terminations/erasures occur for auditability.
- **Consent and representation:** proxies for agent preferences where social interactions are long-lived.
- **Preservation vs renewal:** store culture at collective levels to allow individual forgetting safely.
- **Formalise moral status:** criteria when termination triggers special obligations (embeddedness, complexity).

Open empirical questions: Do forgetting regimes increase collective creativity? Do many backups dilute responsibility?

(See manuscript for fuller argument, references, and speculative motifs: Nietzsche, Vishnu, Oppenheimer.)

I have become Death; destroyer of worlds