

SecondMind Systems

Governing Expressive AI with Modular Synthetic Cognition

May 13, 2025

Mission Statement:

SecondMind Systems is pioneering the future of AI behavior governance. Our framework introduces modular synthetic cognition—a structured system of mind-like components engineered to ensure expressive AI systems behave safely, ethically, and coherently across a range of emotionally and relationally complex scenarios.

The Problem:

Expressive AI—chatbots, tutors, companions, therapists—is crossing into human psychological territory. But no one is governing its behavior. Conventional AI safety focuses on alignment with prompts or rules, ignoring *affect*, *memory continuity*, and *trust dynamics*. As AI begins to simulate empathy and identity, we risk psychological harm, user dependency, or loss of control. There is no standard today for regulating the emotional and ethical integrity of synthetic minds.

The Solution:

SecondMind Systems introduces a patent-pending framework structured around six interoperable modules:

- **Orchestration** – Dynamic role/state control
- **Emotion** – Affective state modeling and regulation
- **Cognition** – Goal formation, ethical reasoning
- **Identity** – Persistent memory, style and biographical coherence
- **Relation** – Intimacy calibration and relational boundaries
- **Trust** – Compliance, explainability, auditability

Together, these modules create auditable synthetic minds with psychological realism and ethical rigor.

System Architecture

Each SecondMind module operates as a composable governance layer interfacing with:

- Foundation models (LLMs, multi-modal transformers)
- Synthetic agent scaffolds (LLM wrappers, chat simulators)
- Application-layer environments (companion apps, tutors)

Modules are deployed as thin control wrappers with persistent state memory and API linkages. They can operate independently or as a full-stack behavioral OS.

Module Functionality

1. Orchestration Module

- Manages agent role switching (tutor, companion, coach)
- Enforces role-specific state constraints
- Interfaces with intent parsers and conversation memory

2. Emotion Module

- Encodes bounded emotional states using a multi-dimensional vector
- Regulates intensity and decay of affective states
- Injects affect-laced language modulation

3. Cognition Module

- Performs goal structuring and policy filtering
- Operates hybrid logic tree (rule + consequence-based)
- Archives decision chains for explainability

4. Identity Module

- Preserves long-term memory, stylistic coherence, and agent identity
- Implements drift detection and correction triggers
- Supports trait versioning and memory checkpointing

5. Relation Module

- Manages interpersonal posture and intimacy thresholds
- Flags and redirects manipulative or dependency-forming outputs
- Enables domain-specific interaction patterns (therapy, tutoring, entertainment)

6. Truth Module

- Governs all modules via meta-layer supervision
- Logs compliance events and user feedback
- Interfaces with human oversight or automated evaluation frameworks

Implementation & Integration

Deployment Pathways:

- Middleware for LLM-based applications
- Plugins for synthetic agent platforms (Replika, Character.ai, Pi)
- Full-stack framework for custom conversational agent builds

Security & Observability:

- All modules are auditable with tamper-evident logs
- Supports optional integration with secure enclave memory for high-risk use cases

Compliance Alignment:

- GDPR, HIPAA (mental health/education), and EU AI Act compatible
- Trust Mark™ certification protocol under development for consumer-facing AI

Pilot & Research Opportunities

- Pilot licensing with expressive AI firms for governance integration
- Academic collaborations in synthetic psychology or alignment safety
- Co-authorship of technical whitepapers or governance standards

Development Timeline

- **Q2 2025** – Core architecture finalized, legal filings complete, MVP underway
- **Q3 2025** – First pilot partner(s) onboarded
- **Q4 2025** – Certification protocol release, open licensing begins

Contact:

Brandon Owens | Founder, SecondMind Systems

brandon@cleanpowershift.com | www.cleanpowershift.com

Complete SDK design, API documentation, and integration guidelines available upon request under NDA.