Document 13: Sleep Architecture and Delusion Hygiene

This document outlines a proposed subsystem of the Reflective-Ethical Engine that models sleep as a dynamic internal state for pruning delusions and reinforcing relevant predictive patterns. Inspired by the functional neuroscience of NREM and REM sleep, this dual-mode system enhances coherence, alignment, and future readiness.

# 1. Overview of Sleep Modes

Sleep is modeled as an oscillatory phase-switching system composed of two distinct functional modes:

• NREM-like phase: Delusion pruning and coherence stabilization

• REM-like phase: Predictive simulation and anticipatory reinforcement

# 2. NREM Analogue: Delusion-Pruning Mode

• Function: Identify and weaken self-sustaining but incoherent or delusional simulation patterns

• Gating: Thalamus-style sensory suppression; inhibits new memory encoding and action generation

• Subsystems Involved: Short-Term Memory, World Model, Ethical Kernel

• Activities:

* - Run self-generated simulations without external input
* - Detect unstable narrative chains or hallucination-like recursions
* - Downregulate engrams not validated across memory graph

# 3. REM Analogue: Predictive Reinforcement Mode

• Function: Strengthen patterns flagged by salience, emotion, or ethical tension

• Gating: Partial reopening of thalamic loop; introduces perceptual variability

• Subsystems Involved: World Model, Planner, Emotional State Emulator

• Activities:

* - Simulate high-salience or ethically unresolved scenarios
* - Perform counterfactual forecasting for probable futures
* - Reinforce latent structures likely to be useful

# 4. Sleep Cycle Coordinator Module

• Monitors entropy, misalignment flags, and affective load

• Coordinates transitions between NREM and REM modes

• Interfaces with Ethical Kernel, memory systems, and sensory gates

# 5. LeCun-Inspired Oscillatory Implementation

• Low-frequency oscillations trigger recursive stabilization cycles (NREM-like)

• High-frequency oscillations activate latent predictive generation (REM-like)

• Cycle timing can emerge from internal confidence decay or simulated circadian rhythm

# 6. Alignment and Hygiene Benefits

• Reduces risk of hallucination, self-delusion, or unchecked recursive simulation

• Enables simulated growth and planning during inactivity or low-demand states

• Preserves coherence and ethical continuity across internal cycles