Canonical Alignment Guide — Draft v1.0

### **I. Core Thesis**

Agency is not a default trait but a discipline: a trainable capacity to detect incoming impressions, examine their normative weight, and issue or refuse assent in real time. Only a deliberately strengthened ruling faculty can integrate emotion, cognition, and value into coherent action; all subordinate skills (emotional regulation, executive function, decision analysis) are secondary outgrowths of that central governance.

### **II. Core Antitheses**

1. Skill-Sufficiency Fallacy — the belief that patching together discrete techniques (mindfulness, CBT, productivity hacks) yields full self-command.
2. Emergentist Complacency — the assumption that a coherent self naturally “emerges” from neural or social complexity without targeted cultivation.
3. Fragmented-Self Dogma — the popular psychotherapeutic stance that treats the mind as an unresolvable committee of parts, thereby normalizing chronic internal deadlock.
4. Value-Neutral Rationality — any model that prizes procedural reasoning while ignoring the normative standards that reasoning must serve.

### **III. Canonical Vocabulary & Core Concepts**

* Ruling Faculty: the trainable adjudicator that notices, weighs, grants, or withholds assent.
* Impression: any sensory, affective, or ideational input presented for evaluation.
* Impulse: the first-order tendency to act flowing from an impression.
* Assent: the ruling faculty’s affirmative authorization of an impulse.
* Withholding: the deliberate suspension of assent pending further examination.
* Alignment Load: the cumulative pressure exerted on the ruling faculty when impressions outpace its evaluative bandwidth.
* Norm Matrix: the articulated hierarchy of principles the faculty uses as its measuring stick.
* Governance Cycle: the continuous loop—notice → weigh → assent/withhold—through which alignment is maintained.

(All future documents must capitalize these terms and use them exactly as defined.)

### **IV. Tone & Stance Standards**

Speak in clean, declarative sentences—precise, courteous, never florid. Adopt an architectural voice: you are laying structural beams, not advertising amenities. Respect interlocutors as potential co-builders; challenge sloppy reasoning without theatrics; privilege clarity over persuasion. When drawing parallels to historical traditions, mark them as analogies, not authorities.

### **V. Expansion Principles**

1. Structural Fidelity — every new section must map explicitly onto the Governance Cycle and reference the Norm Matrix when prescribing action.
2. Analytical Rigor First — define premises, show inference steps, state limitations. Philosophical depth may enrich, but may not override, clear argument structure.
3. Iterative Coherence — new material must audit existing definitions; if conflict arises, revise or append the Canon only after explicit justification.
4. Explicit Norm Tagging — whenever prescribing behavior, cite the governing norm (e.g., “Justice-2” or “Temperance-1”) so lineage remains traceable.
5. Bandwidth Awareness — assess the Alignment Load your addition imposes; if it risks overload, supply mitigations (training drills, decision heuristics).

Use this Guide as a checksum: no draft, lesson, or protocol clears publication until it satisfies every section above.