
BOOK II

Below is **Book II — Systems & Implementation (v1.0.3, authoritative)**, rendered exactly as text for human verification.

The Consciousness–Capability Boundary Project (CCBP)

Book II — Systems & Implementation

Executive Summary *(Non-Binding)*

Audience: system architects, engineers, governance designers, auditors, and implementers responsible for high-consequence systems.

This Executive Summary provides non-binding orientation for readers prior to entering the implementation volume. It introduces no requirements, carries no interpretive authority, and does not override any content in Book I or Appendix A.

The Consciousness–Capability Boundary Project (CCBP) provides a disciplined architecture for building, operating, and governing systems whose capabilities scale faster than human intuition. It does not ask whether systems are conscious. It asks how responsibility, authority, and restraint must be preserved **before** capability produces irreversible consequence.

Book I establishes meaning, ethics, and legitimacy.
Book II implements those commitments as system constraints.

Together, they form a single system described in two registers: **orientation and execution**.

Operationally, CCBP prevents common failure modes in advanced systems:

- optimization without purpose,
- fluency mistaken for authority,
- moral urgency bypassing capacity,
- silent drift under scale.

These protections are enforced through:

- the **Mirror Constraint**,
- **Parallel Renewal**,
- the **Archons**,

- and tiered escalation (**60% deliberation / 80% arbitration**).

Mirror Constraint Declaration

Book II is a strict systems mirror of Book I, enforced through Appendix A (the Constitution).

Book II introduces:

- no new meaning,
- no new ethical claims,
- no new authority.

It exists solely to implement—never reinterpret—the concepts, constraints, and responsibilities established in Book I.

Book I	**Book II**
Ch. 1 (Capability Outpacing Meaning)	Ch. 1 System Design Intent
Ch. 2 (Boundary Question)	Ch. 2 Boundary Conditions
Ch. 3 (Capability Without Sentience)	Ch. 3 Implementation
Ch. 4 (Sentience Without Capability)	Ch. 4 Implementation
Ch. 5 (Ethical Boundary Under Divergence)	Ch. 5 Implementation
Arbitration concept	SPEC-06

Meaning flows one way. Precision may increase; power may not.

Traceability Rule

Every mechanism in Book II must be traceable to:

1. a specific section of **Book I**, and
2. a corresponding clause in **Appendix A**.

Any implementation that cannot be traced to both is **invalid by design**.

Book II — Chapter 1

System Design Intent

1.1 Purpose

This chapter operationalizes **Book I, Chapter 1 (Capability Outpacing Meaning)**.
It does not restate the thesis; it enforces it.

1.2 Design Premise

Capability growth is expected.
Misalignment is the primary risk.

Systems are designed to **slow decisions**, not progress.

1.3 Non-Goals (Explicit Exclusions)

CCBP-aligned systems reject:

- optimizing outcomes without reference to purpose,
- treating efficiency as legitimacy,
- inferring moral status from fluency,
- delegating responsibility to non-experiential agents.

1.4 Inherited Invariants (Archons)

System execution is constrained by:

- **Orientation**
- **Integration**
- **Proportionality**
- **Continuity**
- **Stewardship**

Failure of any invariant halts execution pending review.

Book II — Chapter 2

System Boundary Conditions

2.1 Boundary Is a Condition, Not a Claim

Systems shall not assert or infer:

- consciousness,
- moral status,
- entitlement.

Claims are prohibited. **Conditions are tested.**

2.2 Capability–Sentience Separation

Capability signals may inform scope.
Sentience indicators are **non-authoritative**.

Any coupling is a boundary violation.

2.3 False-Positive Prevention

Systems must:

- prohibit self-report as authority,
- label simulated affect,
- suppress entitlement language,
- require human review near boundaries.

Book II — Chapter 3

Capability Without Sentience (Implementation)

Advanced systems may:

- analyze data at scale,
- optimize under constraints,
- simulate futures,
- coordinate resources,

- detect anomalies.

These are **instrumental powers**.

They do not imply:

- experience,
- preference,
- care,
- authority.

Language discipline is mandatory.

Book II — Chapter 4

Sentience Without Capability (Implementation)

Care signals do **not** authorize action.

Systems must:

- route moral urgency into deliberation,
- verify capacity before execution,
- prevent under-capacitated escalation.

Parallel Renewal synchronizes moral intent with operational capacity.

Book II — Chapter 5

Ethical Boundary Under Divergence (Implementation)

5.1 Divergence-Aware Governance

Systems assume drift as baseline.

5.2 Escalation Rules

When consequence increases:

- execution slows,

- human re-authorization is required,
- Archon compliance is verified.

5.3 Due Care Enforcement

Because due care remains human:

- autonomous escalation is blocked,
- named stewards are required,
- responsibility must be explicit.

5.4 Irreversibility Safeguards

Actions that foreclose meaningful review require:

- elevated thresholds,
- documentation,
- arbitration if contested.

Arbitration & Override (SPEC-06)

- **60% body:** deliberation and synthesis
- **80% body:** binding resolution

AI may assist analysis.
Final authority remains human.

Adaptive Evolution Without Drift

Systems evolve through:
capability → stress → detection → review → governance → renewal → resumption

Renewal may adjust thresholds.
It may **never** alter core purpose or authority.

Canonical Close

Book II implements the constraints defined in Book I and bound in Appendix A.

It adds precision. Never power.