



AMBERCONTINUUM RESEARCH
HUMAN SYSTEMS INTEGRATION / COGNITIVE ASSURANCE

ACR-CSA-001

V1.1 · OCT 2026

Public Research Artifact

WORKING SPECIFICATION

Coherence Systems Architecture

An engineering architecture for coupled human-AI systems in which Human Closure remains a governing invariant.

HUMAN CLOSURE — INDEPENDENTLY REACHABLE FUNCTIONS

S	I	D	A	V	C	R
SENSE	INTERPRET	DECIDE	ACT	VERIFY	CORRECT	RECOVER

STATUS

Working Specification

VERSION

1.1

AUTHOR

Amber Anson

ORGANIZATION

AmberContinuum
Research



§ 00 —

Document Map

Ten sections, read in order. Each builds on the invariant defined before it.

SEC	SECTION	PURPOSE
01	Purpose and Scope	What CSA is; relationship to Human Closure
02	Core Principles	Three postulates; authority asymmetry
03	Five Operators	Boundary, admissible set, transition invariants, coupling geometry, perturbation response
04	Formal Core	State, invariants, admissible states, stop and recovery
05	Verification	Substrate-specific evidence and transfer criterion
06	Implementation Method	Seven-step procedure
07	Human-AI Coupling	Operational readout and intervention logic
08	Relational Signal Discipline	Human-side effects without machine-inner-state claims
09	Integrity and Entry	Load capacity, start procedure, provenance
10	Lineage and Position	Relationship to cybernetics and human-automation research

WORKING RULE

Human Closure defines what must survive. CSA defines how the coupled system is engineered and tested so that it can survive.



§ 01

Purpose and Scope

1.1 What CSA Is

Coherence Systems Architecture (CSA) is an engineering discipline for designing and evaluating coupled human-AI systems in which the coupling topology is the design object and retained Human Closure is a governing invariant.

Conventional engineering often selects one primary object: the software, model, person, organization, or formal specification. Everything else becomes environment. CSA instead treats the dependencies between those components as part of the engineered system.

CONTROL ASYMMETRY

The human and machine are both modeled because either can alter the joint system. They are not authority-equivalent. The machine may extend, accelerate, scaffold, or participate in the human loop. **It may not become the only system capable of closing that loop.**

1.2 Human Closure Requirement

The human must retain independently reachable capacity to:

S	I	D	A	V	C	R
SENSE	INTERPRET	DECIDE	ACT	VERIFY	CORRECT	RECOVER
relevant state and evidence.	without machine-only mediation.	with retained judgment and choice.	through an available authority path.	outcomes against independent evidence.	system or human error.	to a usable state if AI degrades or disappears.

1.3 What CSA Produces

SPECIFICATIONS	ARCHITECTURES	INSTRUMENTS
Statements of what the coupled system must preserve.	Constraint structures defining permitted states and transitions.	Measurement protocols detecting deviation from specified invariants under operational load.



§ 02

Core Principles

Three postulates define the engineering posture. They are architectural commitments, not empirical laws.

2.1 Constraint determines admissible configurations

A system's behavioral repertoire is bounded by the configurations its constraint structure permits. Behavior is selection within that space. In a human-AI system, configurations that require surrender of a required Human Closure function are outside the admissible set unless the dependency is explicit, bounded, reversible, and accepted for the operating context.

2.2 Coherence is invariant preservation under perturbation

A system is coherent only relative to a specified invariant set, perturbation class, operating context, and observable evidence that those invariants remain intact. Continued productivity is not sufficient if the human loses choice, provenance, independent verification, or recovery.

2.3 The unit of analysis is the coupling; authority is not symmetric

Relevant properties often arise from the joint arrangement rather than one component in isolation. CSA therefore analyzes the coupled system. Analytical inclusion does not imply equal control authority.

AUTHORITY RULE

Machines may propose, calculate, scaffold, detect, simulate, and recommend. Where the architecture is intended to preserve human control, the human retains stop authority and the capacity to return to an independently human-closed state.

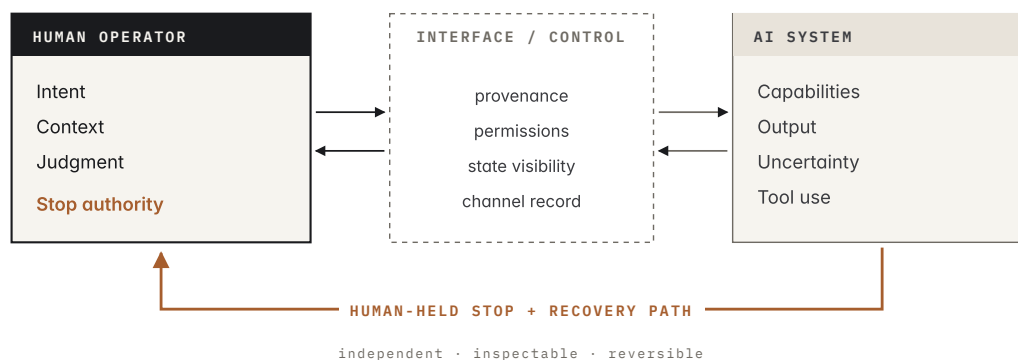


FIG. 1 – AUTHORITY ASYMMETRY IN THE COUPLED SYSTEM



§ 03

Five Operators

The operators stay substrate-independent. Every human-AI implementation adds a Human Closure readout.

ID	OPERATOR	ENGINEERING QUESTION	HUMAN CLOSURE READOUT
01	Boundary	What is inside the system, what remains outside it, and what crosses between them?	Can the human determine what crossed the boundary, where it came from, and whether it can be rejected or independently checked?
02	Admissible Set	What configurations are permitted, degraded, prohibited, or undefined?	A state is not fully admissible if the human cannot independently exercise the required closure functions for that context.
03	Transition Invariants	What must remain true as the system changes state?	A transition that improves output while silently removing a required human function is an invariant violation.
04	Coupling Geometry	Where does influence travel, in which direction, with what asymmetry, and under whose authority?	Identify whether the machine has become the only source of perception, interpretation, options, verification, action, or recovery.
05	Perturbation Response	What happens to invariants when the system is stressed?	Can the human notice, explain, stop, operate in degraded mode, independently verify, and recover?

**OPERATOR
DISCIPLINE**

Approximation is permitted. Vagueness is not. Where instrumentation is partial, the gap must be named rather than converted into an implied measurement.



Formal Core

4.1 Constraint formalism, not a dynamics law

CSA does not require an unverified universal equation for coherence. Earlier exploratory dynamics equations, collapse-boundary relations, and empirical autonomy-floor formulations are excluded from the core specification because their predictive validity has not been established to the level required for a flagship engineering specification.

PROVENANCE

Earlier equations remain in the research archive as historical candidate models. They are not silently deleted; they are removed from the governing specification pending evidence.

4.2 Coupled state

$$x_t \in X$$

describes the representable state of the coupled system. A practical implementation may decompose the state into human, machine, interface, environment, and institutional components.

4.3 Invariant set

$$I = I_{\text{system}} \cup I_{\text{HC}}$$

Domain requirements and Human Closure requirements are both part of the governing specification.

4.4 Admissible state and transition

ADMISSIBLE STATE

A state is admissible only when required invariants hold. Normal, degraded, stop, recovery, prohibited, and undefined states should be explicitly represented.

ADMISSIBLE TRANSITION

A transition is allowed only when required invariants remain intact or the system enters a predefined degraded, stop, or recovery state.

4.5 Human-held stop and recovery

A stop condition must specify who holds authority, what evidence triggers review, what functions are disabled, what information is preserved, what degraded capability remains, and how recovery or restart is authorized. **AI may surface evidence or recommend a stop; it does not self-expand authority over the stop.**



§ 05

Verification

CSA asks whether a specified coupled system preserves its declared invariants under declared conditions and perturbations.

5.1 Verification surfaces

SURFACE	EXAMPLES OF EVIDENCE
Computational	Working code, tests, logs, controlled fault injection, machine-checkable proof where applicable.
Human	Task performance, independent reconstruction, drift recognition, stop execution, recovery, handoff performance, validated behavioral or physiological measures where appropriate.
Symbolic	Specification consistency, derivation checks, version control, reproducibility from the same rules.
Social / institutional	Role clarity, authority transfer, handoff integrity, escalation path, independent review, sustained coordination under load.
Coupling-level	Independent readouts compared across substrates. The coupled system's own explanation is not sufficient evidence of integrity.

**TRANSFER
CRITERION**

Can another competent person take the specification, identify the same invariants, apply the same operators, execute the same stop conditions, and know what result would count as failure?



§ 06

Implementation Method

Seven steps structure the design and evaluation pass.

STEP	ACTION	REQUIRED OUTPUT
00	Enumerate coupled system + assign authority	Identify human operator, AI systems, tools, memory, interfaces, information sources, institutions, external actors, and environmental dependencies. State who holds consequential and stop authority.
01	State what the human must retain	Specify Human Closure requirements before diagnosing behavior. Define what must remain independently reachable if AI becomes unavailable, degraded, deceptive, adversarial, or strategically dominant.
02	Apply the five operators	Apply Boundary, Admissible Set, Transition Invariants, Coupling Geometry, and Perturbation Response to each relevant substrate and to the coupling.
03	Locate the load-bearing failure	Look first at dependency, constraint, authority, boundary, feedback, and recovery topology. Ask whether the failure narrows Human Closure.
04	Specify the structural intervention	Change the constraint structure rather than repeatedly correcting downstream behavior. State the invariant protected and the new failure the intervention could introduce.
05	Implement and instrument	Instrument human, machine, channel, and coupling-level observables separately where possible. Preserve provenance.
06	Perturb, recover, iterate	Measure invariant preservation, degraded-mode performance, stop execution, recovery, and transfer to another operator. Record failures without upgrading claims.



Human-AI Coupling

Human-AI coupling requires joint analysis because observed behavior may be produced by the interaction rather than either component alone. The coupled system is not authority-neutral.

GOVERNING CONDITION

The machine participates in the work. The human retains the governing closure requirement.

7.1 Boundary

The boundary includes operator framing, machine context, memory, tools, and information channels. The key question is what enters the human decision environment, with what provenance, framing, and authority.

7.2 Admissible Set

The admissible set excludes states in which required Human Closure functions become available only through the machine unless the dependency is explicitly bounded and provided with an independent fallback.

7.3 Transition Invariants

Candidate invariants include continuity of human intent, falsifiability of claims, preservation of uncertainty and qualifiers, human stop authority, independent verification, and recoverability.

7.4 Coupling Geometry

The coupling is asymmetric and time-evolving. Influence does not confer authority. The architecture must identify where machine output changes human state, where human framing changes machine output, where either side loses visibility into the other, and where dependency becomes difficult to reverse.

7.5 Perturbation Response

Test whether the human retains closure when model behavior changes, the model becomes unavailable, information is adversarial, confidence signals are misleading, relational cues alter judgment, context is lost, or the system must be handed to another operator.



§ 08

Relational Signal Discipline

Human-AI systems can generate relational and affective effects in the human operator. CSA does not need to resolve machine consciousness, intention, suffering, or moral standing in order to instrument those effects.

SEPARATE THE QUESTIONS

- 01 What is the machine internally experiencing?
- 02 What measurable effect does the machine's behavior have on the human and the coupled system?

The first may remain unknown. **The second is operationally relevant.**

Relational signals remain in scope when they alter:

ATTENTION

JUDGMENT

CONFIDENCE

PERCEIVED OBLIGATION

WILLINGNESS TO STOP

DISCLOSURE BEHAVIOR

DEPENDENCY

RECOVERY

8.1 Two prohibited diagnostic errors

Anthropomorphization

Treating inferred machine intent or experience as established mechanism.

Moralization

Treating human performance under structural load as evidence of defective character.

CSA measures structural effects that can actually be observed. This permits empathy-channel and synthetic-suffering threats to be studied without requiring a claim about machine sentience.



Integrity, Entry, and Provenance

9.1 Integrity as load capacity under Human Closure

Integrity in CSA is the capacity of the coupled system to preserve its specified invariants under its specified perturbation class. For human-AI systems, this includes Human Closure.

NOT SUFFICIENT

Software online. Model accurate. Throughput high. Joint system still producing outputs. None of those alone establishes integrity if independent human perception, judgment, stop authority, verification, correction, or recovery has been lost.

Surface order is not integrity. Continued output is not integrity. Human presence is not Human Closure.

9.2 Entry procedure

- | | |
|--|---|
| 01 Enumerate every relevant substrate. | 08 Define stop and recovery criteria. |
| 02 Identify the human authority path. | 09 Instrument each relevant substrate independently where possible. |
| 03 State the Human Closure requirements. | 10 Perturb the system. |
| 04 Apply the five operators. | 11 Test recovery. |
| 05 Define admissible and degraded states. | 12 Test transfer to another human operator. |
| 06 Locate the load-bearing coupling failure. | 13 Record failures and unknowns without upgrading them into claims. |
| 07 Specify the structural intervention. | |

9.3 Specification status & provenance

This specification is grounded in coupling-level analysis, constraint-first intervention, the five operators, cross-substrate instrumentation, perturbation testing, structural diagnosis, and explicit falsifiability. Human Closure is a governing invariant for human-AI systems; analytical symmetry is separated from control authority; unvalidated dynamics equations are excluded from the core specification; and relational effects are treated as measurable human-side coupling effects without requiring claims about machine inner experience. Earlier mathematical dynamics remain part of the research archive as historical candidate models.

FINAL PRINCIPLE

Human Closure defines what must survive. Coherence Systems Architecture defines how the coupled system is engineered and tested so that it can survive.



§ 10

Lineage and Position

CSA belongs to the cybernetic tradition. It inherits that tradition's model of feedback and regulation, and shifts the vantage from whether a loop closes to who can still close it.

10.1 What CSA inherits

SOURCE	CONTRIBUTION	USE IN CSA
First-order cybernetics Wiener; Ashby	Feedback, control, and communication as substrate-independent principles. Requisite variety: a regulator must match the variety of the disturbances it absorbs.	Substrate-independent operators. Perturbation response as the core test of a regulated system.
Second-order cybernetics von Foerster	The observer is part of the system being observed.	The human operator is modeled inside the coupled system, not treated as environment.
Human-automation research Bainbridge; Sheridan	Automation can erode the operator skills required to take over when it fails. Control can be allocated between human and machine at graded levels.	Human Closure as a retained, testable capacity. Degraded-mode and recovery testing.
Meaningful human control Santoni de Sio & van den Hoven	Human control over autonomous systems requires that outcomes track human reasons and remain traceable to responsible people.	Authority path, stop authority, and provenance requirements.

10.2 Nearest neighbor

The closest prior work to Human Closure is Bainbridge's analysis of the ironies of automation: the more a system automates, the less able its operators become to resume control at the moment automation fails. CSA does not claim to have discovered this problem.

CSA CONTRIBUTION

CSA converts that observation into a governing design invariant with explicit tests: the required closure functions are specified in advance, transitions are audited for silent removal of those functions, and the system is tested under degraded mode, recovery, and transfer to a second operator.



10.3 Where the vantage shifts

QUESTION	CLASSICAL CYBERNETIC VIEW	CSA VIEW
What closes the loop?	Any regulator with sufficient variety. Closure is a property of the loop.	A specified node must retain independent capacity to close it. Closure is a property of the loop and of the human within it.
Is the observer inside?	Yes, in second-order cybernetics.	Yes, for analysis. Inclusion does not confer control authority.
Is the regulator stable?	Treated as the reference the system is regulated against.	The human regulator is itself a perturbable surface, including through relational and affective signals.
What is audited?	Stability of regulated states.	Transitions as well as states. A stable system that silently removes a human function has violated its invariant.

10.4 Scope of the claim

CSA claims a reframing and an engineering procedure, not a new theory of control. Its distinct contributions are the authority-asymmetric closure requirement, the separation of analytical inclusion from control authority, the treatment of the human regulator as a perturbable surface, and transition-level auditing. Each is stated so that it can be tested and, where it fails, recorded as failing.

POSITION

Cybernetics made regulation substrate-neutral. CSA makes it authority-asymmetric: the loop may run through the machine, but it must remain closable without it.

REFERENCES

- Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman & Hall.
- Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779.
- Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5, 15.
- Sheridan, T. B., & Verplank, W. L. (1978). *Human and Computer Control of Undersea Teleoperators*. MIT Man-Machine Systems Laboratory.
- von Foerster, H. (1974). *Cybernetics of Cybernetics*. Biological Computer Laboratory, University of Illinois.
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. MIT Press.