[ CAPABILITY PROFILE · AUTHREX-06 ]

# CARA · CONTROLLED AUTHORITY RECOVERY ARCHITECTURE

Make recovery from authority lockout structured, deterministic, and impossible to short-circuit, because turning the system off is not always a safe or available option.

**TESTED** MODELED

Name variants: Patents-page expansion: Control Authority Regulation Architecture; Zenodo record 18917790 title: Cognitive Authority Recovery Architecture. Same framework; original artifact titles preserved. GREP phases on the CARA page: Guard, Reduce, Evaluate, Promote (the BLADE-AV page uses a Restrict and Execute phrasing, not carried).

| | |
|---|---|
| PRIMARY FUNCTION | Deterministic recovery from lockout through GREP phases (Guard, Reduce, Evaluate, Promote) with a terminal non-compensatory policy gate, so that a single favorable signal cannot restore authority; returns control to SATA, closing the loop. |
| MISSION AREAS | Spacecraft, uncrewed maritime and undersea systems, and industrial control systems where shutdown is not a safe state; any platform that must recover from A0 in a governed way. |
| EVIDENCE ANCHORS | Zenodo DOI 10.5281/zenodo.18917790; self-run 10-million-iteration enumeration and 68-state control-flow analysis (not independently reproduced); seeded browser simulation; formal-coverage record: CARA.tla surrogate countermodel with a published strong-form counterexample. |
| CURRENT STANDING | Demonstrated in simulation; enumeration artifacts are self-run; the formal record shows a replayed re-arm token defeats the naive form; requirement stated (token bound to a nonce); not independently reproduced. |
| SELF-ASSESSED TRL | TRL 3. No formal TRL determination has been performed. |
| INTELLECTUAL PROPERTY | U.S. provisional application 64/000,170, filed March 9, 2026 (receipt 74767602), reported as filed; unexamined; confers no enforceable rights. |

[ MISSION PROBLEM ]

Recovery is where most authority architectures are silent. After a lockout, the practical pressure is to restore autonomy as soon as any signal looks good, and a recovery path that can be short-circuited by one favorable reading, or replayed by an attacker who captured a previous re-arm, becomes the widest hole in the system. For platforms where shutdown is itself hazardous (a spacecraft in a bad attitude, an undersea vehicle at depth, a process plant mid-cycle) the consequence of unstructured recovery is either a premature return to full authority or an authority cliff with no way back.

[ OPERATIONAL RELEVANCE ]

- Spacecraft safe-attitude hold and staged recovery on refreshed evidence (BLADE-SPACE, SPACECYBER reference designs).
- Uncrewed maritime and undersea vehicles where link loss must produce governed fallback rather than silence.
- Industrial control and utility operations where a controlled return to automatic mode must not be triggered by a single sensor.
- Ground robotic platforms recovering from an A0 (revoked) safe-stop.

[ CAPABILITY DESCRIPTION ]

CARA activates when HMAA authority enters A0 (revoked). Guard is an immediate safe-stop with all autonomous commands disabled; Reduce permits minimal safe behaviours only (return-safe or hover-hold on trusted sensors); Evaluate monitors trust recovery and requires sustained improvement before any restoration; Promote restores authority gradually, with constrained operation until full trust is confirmed. The recovery behaviours run safe-stop, return-safe, hover-hold, crawl-mode, degraded-teleop. The terminal gate is non-compensatory: every required condition must be independently satisfied, and a strong signal on one condition never offsets a failed one. Recovery returns control to SATA rather than directly to the controller, which closes the loop and means promotion is subject to the same trust assessment as initial authority. The formal record adds a requirement: the re-arm token must be bound to a nonce, because a replayed token satisfied the bookkeeping while ground truth recorded that it was never issued.

[ CARA · ARCHITECTURE, ASSURANCE, EVIDENCE ]

[ TECHNICAL ARCHITECTURE ]

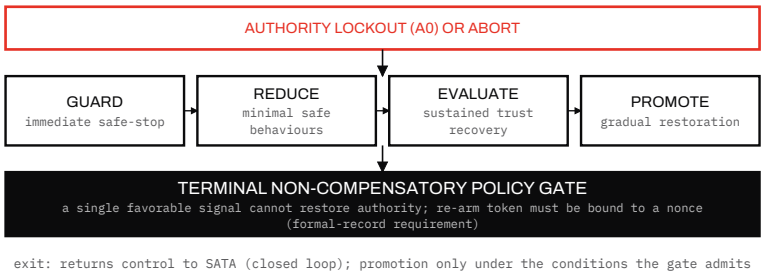


Figure 1. CARA technical architecture. Encoding: solid = implemented; dashed = specified; gray = external; filled = enforcement point.

[ HOW IT WORKS ]

1. Enter Guard on A0 (revoked): immediate safe-stop, all autonomous commands disabled, lockout cause recorded.
2. Reduce: minimal safe behaviours only (return-safe or hover-hold on trusted sensors); shed residual authority and pending actions.
3. Evaluate: monitor trust recovery through SATA and re-run any policy-specific conditions; each must pass independently and improvement must be sustained.
4. Apply the terminal non-compensatory gate; if any condition fails, remain in Guard or Reduce.
5. Promote: restore authority one tier at a time through HMAA, subject to hysteresis; a re-arm token bound to a nonce authorizes each promotion.
6. Return control to SATA so that the restored authority is continuously re-assessed.
7. Record each phase transition, the gate evaluation, and the token to the ledger.

[ TECHNICAL DIFFERENTIATION ]

A kill switch terminates and leaves recovery to procedure; conventional fail-safe designs specify the fall but not the climb. CARA makes recovery a first-class system function with its own gate, so a single favorable signal cannot restore authority and a replayed authorization cannot re-arm the system. The differentiation is structural; the enumeration evidence is self-run and no comparative measurement against another recovery scheme is claimed.

[ ASSURANCE MECHANISMS: DESIGN PROPERTY VERSUS ESTABLISHED PROPERTY ]

| Mechanism | Status |
|---|---|
| Non-compensatory terminal gate | Design property; exercised by the self-run enumeration and simulation |
| Deterministic phase sequence | Design property; 68-state control-flow analysis (self-run) |
| No promotion on a single favorable signal | Design property; exercised in simulation |
| Re-arm token bound to a nonce | REQUIREMENT from the published counterexample; not established for the shipped engine |
| Return to SATA (closed loop) | Implemented in the seeded consoles |

[ TECHNICAL EVIDENCE ]

| Evidence | Result | Source / artifact |
|---|---|---|
| Enumeration of the recovery state space | 10,000,000 iterations, self-run; 68-state control-flow enumeration | Zenodo 10.5281/zenodo.18917790 (site claim; not independently reproduced) |
| Seeded browser simulation | GREP path exercised under injected fault clearance (scenario S7 in the evaluation protocol) | burakoktenli.com/cara-simulation |
| Strong-form counterexample | Replayed re-arm token: bookkeeping records a token present, ground truth records it never issued | authrex.systems/formal-coverage |
| Surrogate countermodel | CARA.tla: finding about the mechanism class | formal-coverage record |

[ EVIDENCE SNAPSHOT ]

| EVIDENCE SNAPSHOT | |
|---|---|
| Implemented | Yes (simulation engine) |
| Executable artifact | Yes |
| Formal model | Partial (surrogate) |
| Simulation evidence | Yes |
| Hardware validation | No |
| Field validation | No |
| Independent validation | No |
| Government evaluation | No |
| Operational deployment | No |

[ LIMITATIONS ]

- The 10-million-iteration enumeration is a self-run site claim; no independent reproduction exists.
- The nonce-bound re-arm token is a stated requirement, not a shipped implementation.
- Recovery timing (mean recovery time metric) has only been measured in idealized simulation.
- Simulation and formal-model evidence only; all data synthetic; no hardware-in-the-loop test, no field validation, no independent replication, no deployment.
- Not independently reproduced or technically verified by another party; second-environment reruns are confirmation by the same producer.

[ TECHNOLOGY READINESS (SELF-ASSESSED; NO FORMAL TRA) ]

| | |
|---|---|
| CURRENT ASSESSED TRL | TRL 3 (self-assessed against the DoD nine-level scale; no formal technology readiness assessment has been performed). |
| BASIS | Recovery logic implemented and exercised in simulation with self-run enumeration; no laboratory validation on a representative platform. |
| EVIDENCE SUPPORTING | concept and phases specified; simulation implementation; self-run enumeration |
| EVIDENCE REQUIRED FOR NEXT TRL | (1) Independent reproduction of the enumeration artifacts. (2) Implementation and check of the nonce-bound token requirement. (3) Recovery-path trial on a software-in-the-loop platform with injected lockout and fault clearance (Phase 4). |

[ INTEGRATION PROFILE ]

| | |
|---|---|
| INPUTS | ABORT or lockout event with cause; policy conditions for promotion; SATA re-assessment; re-arm authorization. |
| OUTPUTS | Phase state; gate verdict; promotion requests to HMAA; ledger entry. |
| INTERFACES | HMAA (A0 entry, promotion), SATA (re-assessment), platform safe-state interface (external interlock). |
| DEPLOYMENT | Software module; browser console for evaluation. |
| SECURITY BOUNDARY | Inside: phase logic, gate. Outside: safe-state actuation, token issuer. A replayable token defeats the gate until the nonce requirement ships. |
| DATA REQUIREMENTS | Synthetic; a government-defined safe-state and recovery policy for a named platform would enable Phase 2. |

[ TRANSITION ROADMAP ]

| Phase | Activity for this capability | Who can execute it | Exit evidence |
|---|---|---|---|
| 1 Technical review | Review of the artifacts listed in the evidence table | Program alone, with a government reviewer | Reviewer findings; correction log entries |
| 2 Government-defined simulation | Sponsor scenarios on the seeded console | Program with sponsor scenarios | Deterministic, replayable results; hash-chained records |
| 3 Degradation and failure injection | Evidence lapse, spoofing, delayed human response | Program with sponsor parameters | False-restriction, miss, and contraction-latency figures |
| 4 Government testbed / SIL | Integration with a representative controller or interface | Requires sponsor interface and testbed | SIL integration report |
| 5 Hardware-in-the-loop | Representative interference; timing on real hardware | Requires sponsor laboratory or rig | Measured latencies; TRL 4 to 5 evidence |
| 6 Relevant-environment demonstration | Independent evaluation and red team | Requires independent evaluator | Independent report |
| 7 Program integration | Only after earlier gates succeed | Requires transition partner | Not claimed to be available |

[ GOVERNMENT ENGAGEMENT REQUEST (SPECIFIC) ]

- Recovery-path review by a platform safety authority (spacecraft, maritime, or ICS).
- A government-defined recovery policy (conditions for promotion) for a named platform class.
- Independent reproduction of the enumeration artifacts from the deposit.

[ REVIEWER QUESTIONS THIS PROFILE ANSWERS ]

What is it: see capability description. Why does DoD care: see operational relevance. Where would it fit: see integration profile. Does it exist: see evidence snapshot. What evidence supports it: see technical evidence. Has anyone independent tested it: no. Has it been tested in the real world: no. What does it rely on: see limitations and security boundary. What can cause it to fail and what does it do then: see assurance mechanisms and the published counterexamples. How mature: see the TRL block. What would the government need to provide: see the request above. Next measurable milestone: first item under evidence required for next TRL.