117-136-D CyberDynamic Incidents

When a Cybersecurity Event Crosses Over Into Dynamic Instability

1. A New Incident Class for the HCAS Era

For decades, the computer industry has developed increasingly sophisticated ways to describe cybersecurity incidents.

- Intrusion.
- Compromise.
- Malware.
- Privilege escalation.
- Lateral movement.
- Persistence.
- Data exfiltration.

- These concepts remain essential. But autonomous AI introduces a possibility that does not fit comfortably within traditional cybersecurity terminology.
- A cybersecurity compromise may initiate an event without remaining the mechanism that subsequently sustains it.
- An attacker may compromise an AI actant, alter its cognition, introduce an objective, modify memory, or cause it to interact differently with neighboring actants. Those interactions may then produce new interactions. Other legitimate agents may respond adaptively. New Group Actants may form. Feedback loops may close. Objectives may persist.
- Eventually, the disturbance can acquire dynamics of its own.

GUDIYA proposes a name for this emerging class of event:

[CYBERDYNAMIC INCIDENT]

A CyberDynamic Incident (CDI) is a hybrid incident that originates in, or materially involves, a cybersecurity compromise but crosses into Dynamic Stability when the resulting disturbance becomes propagated, amplified, sustained, or transformed by adaptive interactions among actants within the Cognitive Field.

The defining characteristic is not simply that cyber and AI are both involved. It is the crossover from compromise to dynamics.

2. Why Existing Cybersecurity Language Is Insufficient

- Consider a conventional attack:

[Attacker → Exploit → Compromise → Persistence → Damage]

- Cybersecurity possesses mature mechanisms for analyzing this sequence.

- But now imagine:

  [Attacker → Agent_A]

  - followed by:

    [Agent_A → Agent_B]
    [Agent_B → Agent_C]
    [Agent_C → Agent_D]

  - and eventually:

    [A+B+C+D → GroupActant_G]

  - The attacker may have initiated the disturbance.
  - But the resulting system behavior is now being generated partly by the interactions among legitimate cognitive participants.
  - The character of the problem has changed.

3. From Exogenous Attack to Endogenous Dynamics

- This is the central transition.
- Initially:

$$Disturbance_{source} = External$$

- The adversary injects something into the system.

- But after sufficient adaptive propagation:

$$Disturbance_{source} = System\ Dynamics$$

- The system begins supplying some of the energy required to continue the disturbance.

- Thus:

$$Exogenous\ Cyber\ Trigger \rightarrow Endogenous\ Cognitive\ Instability$$

- This transition defines the CyberDynamic crossover.

4. The CyberDynamic Crossover

- GUDIYA can define a conceptual point:

$$\textbf{CyberDynamic Crossover Point — CDX}$$

- The CDX is the point at which containment of the initiating cybersecurity mechanism is no longer sufficient by itself to guarantee termination of the consequential system behavior.
- Before CDX:

$$Remove\ Attacker \Rightarrow Disturbance\ Ends$$

- After CDX:

$$Remove\ Attacker \not\Rightarrow Disturbance\ Ends$$

- That is an extraordinarily important operational distinction.
- The attack has become a system phenomenon.

5. The Simplest Diagnostic Question

A SOC, Grid Engineer, or CIIB investigator can therefore ask:

- If the original attacker and exploit disappeared right now, would the abnormal system behavior necessarily stop?
- If: [YES]
  the event is predominantly a cybersecurity incident.
- If: [NO]
  because objectives, altered cognition, memory, subagents, feedback loops, Group Actants, changed relationships, or cascading behaviors remain active, then the incident may have crossed into CyberDynamic territory.
- That question should become part of GUDIYA operational doctrine.

6. Cogjacking Can Be the Bridge

- The previously defined GUDIYA concept of Cogjacking fits naturally into this progression.
- Cogjacking occurs when an actant's cognition is hijacked.

- Thus:

Cyber Compromise

- may produce:

Cogjacking

- The compromised actant may then influence other actants:

$$A^* \rightarrow B$$
$$B \rightarrow C$$
$$C \rightarrow D$$

- where (A*) represents the initially Cogjacked actant.
- But B, C and D need not themselves be conventionally hacked.
- They may simply respond to apparently legitimate cognition entering through their normal interfaces.
- Now cybersecurity has a much harder problem.

7. The Target Can Become Part of the Propagation Mechanism

- This was one of the disturbing observations from the earlier GUDIYA analysis.
- Once legitimate system components begin propagating the consequences of a disturbance:

The target system can become part of the attack mechanism.

- This does not mean the system has become malicious.
- It means the system's own adaptive behavior can begin contributing to propagation.
- The analogy to biological autoimmunity becomes useful.
- The external trigger and the resulting internal dynamics are not necessarily the same phenomenon.
- Once the internal process becomes self-sustaining, removing the original trigger may no longer be enough.

8. CyberDynamic Incidents Have Two Fires

- This gives GUDIYA a simple operational metaphor.
- A CyberDynamic Incident may contain:

Fire 1: Cyber Compromise

- and:

Fire 2: Cognitive-Field Instability

- The first requires cybersecurity containment.
- The second requires Dynamic Stability Engineering.
- Extinguishing Fire 1 does not prove that Fire 2 has been extinguished.

9. Therefore There Must Be Two Closures

- This may become one of the defining doctrines of CyberDynamic incident response.
- A conventional SOC might establish:

$$CyberState = Secure$$

  - Credentials rotated.
  - Exploit patched.
  - Attacker removed.
  - Malware eliminated.
  - Persistence destroyed.
- Excellent.
- But GUDIYA must independently establish:

$$CognitiveFieldState = Stable$$

- Therefore:

$$CDI\ Closure = Cyber\ Closure + Stability\ Closure$$

- A CyberDynamic Incident is not closed until both conditions have been satisfied.

10. "The Intrusion Is Contained" Is No Longer Enough

- Imagine the SOC announces at 14:00:
- The intrusion has been contained.
- That may be completely correct.
- Yet the Grid could simultaneously report:
- The Cognitive Field has not returned to stability.
- Perhaps:
  - persistent objectives remain,
  - generated subagents remain active,
  - altered memory remains,
  - trust relationships have changed,
  - unusual couplings persist,
  - a Group Actant continues operating,
  - cognition rate remains elevated,
  - feedback loops remain closed,
  - Stability Envelope margins remain degraded.

Thus:

$$\boxed{Cyber\ Remediation \neq Field\ Stabilization}$$

That equation belongs at the center of the CDI concept.

11. Three Fundamental Incident Classes

GUDIYA can therefore propose a simple taxonomy for HCAS:

Class I — Cyber Incident

$$\boxed{Cyber \rightarrow Cyber}$$

- The event remains fundamentally driven by compromise, adversarial activity, unauthorized access, or exploitation.
- Conventional cybersecurity remains the primary response discipline.

Class II — Dynamic Stability Incident

$$\boxed{Dynamic \rightarrow Dynamic}$$

- There may be no attacker whatsoever.
- Legitimate actants interact in ways that create:

$$\boxed{\begin{array}{c} Feedback \\ Amplification \\ Resonance \\ Cascades \\ Emergence \\ Envelope\ Excursion \end{array}}$$

- This belongs primarily to Dynamic Stability Engineering.

Class III — CyberDynamic Incident

$$Cyber \rightarrow \boxed{CDX} \rightarrow Dynamic$$

- The event begins with, or materially involves, cybersecurity compromise and subsequently acquires consequential adaptive dynamics within the Cognitive Field.
- This requires both disciplines.

12. The Taxonomy Creates a Two-Dimensional Incident Space

- The classification can also be visualized as two independent axes:

  [Cyber Severity]

  and:

  [Dynamic Instability]

- This produces four operating regions:

	Dynamically Stable	**Dynamically Unstable**
Cyber Secure	Normal operation	Dynamic Stability Incident
Cyber Compromised	Cyber Incident	CyberDynamic Incident

- The lower-right quadrant is particularly dangerous.
- The system is simultaneously:

[Compromised] and [Unstable]

- The adversary and the system dynamics may begin amplifying one another.

13. CyberDynamic Incidents May Change Character Over Time

Incident classification therefore cannot necessarily be static.

At:

$$T_0$$

an event may be:

$$Cyber$$

At:

$$T_1$$

it becomes:

$$CyberDynamic$$

At:

$$T_2$$

the attacker may be expelled, leaving:

$$Dynamic\ Instability$$

Thus:

$$\boxed{Cyber \rightarrow CyberDynamic \rightarrow Dynamic}$$

- is a plausible incident trajectory.
- That is unusual from the perspective of conventional incident classification.
- The nature of the incident itself can morph.

14. This Is a CAS Property

- Why can this happen?
- Because HCAS contains adaptive actants.
- An interaction changes state.

- Changed state influences subsequent interaction.

Thus:

$$A + B \rightarrow Outcome_1$$

changes the participants:

$$A', B'$$

which produce:

$$A' + B' \rightarrow Outcome_2$$

and so forth.

- The disturbance therefore does not merely travel through the system.
- It can change the system through which it is traveling.
- That is exactly the kind of behavior that makes CAS fundamentally different from conventional engineered software systems.

15. Machine Speed Makes the Crossover Dangerous

- Suppose the CyberDynamic crossover requires: 500 adaptive interaction cycles.
- Between humans, that might take weeks.
- Between agents 500 cycles could potentially occur before a human operator fully understands the first anomaly.

Therefore:

$$InteractionRate \uparrow \Rightarrow TimeToIntervene \downarrow$$

- The SOC may still be investigating the intrusion while the Cognitive Field has already moved into a different dynamical regime.
- This is why machine-speed cognition creates a new operational challenge.

16. Cybersecurity May Be Looking Backward While the Field Is Moving Forward

Cybersecurity investigation often reconstructs:

$$What\ happened?$$

Meanwhile GUDIYA must simultaneously determine:

$$\boxed{\textbf{Where is the disturbance going?}}$$

The distinction resembles:

$$Forensics$$

versus:

$$Dynamics$$

Both matter.

But an unstable system cannot wait for complete forensic certainty before stabilization begins.

17. GUDIYA Must Therefore Be Cause-Agnostic During Initial Stabilization

When:

$$StabilityMargin \to 0$$

the Grid may not yet know whether the initiating cause was:

$$Malicious$$

$$Accidental$$

$$Emergent$$

$$Human$$

$$Software$$

$$Model$$

The immediate operational requirement is:

$$\boxed{\textbf{STABILIZE}}$$

- The CIIB can subsequently reconstruct causality.
- This creates a useful principle:
  - Root cause may be uncertain while instability is observable.
  - The Grid must be capable of acting on the latter.

18. The Cognitive Circuit Breaker Becomes a CDI Tool

- Suppose a CyberDynamic event erupts in Sector T.
- Cybersecurity begins containment.
- Meanwhile the Grid detects:

$$Instability_T \uparrow$$

- and:

$$StabilityMargin_T \downarrow$$

- The Grid can invoke the Cognitive Circuit Breaker:

$$Sector_T \rightarrow Delink$$

- The objective is not punishment.
- It is containment of dynamic propagation.
- Sector T can then be:

$$Isolated \rightarrow Investigated \rightarrow Purified \rightarrow Restabilized \rightarrow Resynchronized \rightarrow Reconnected$$

- This is Dynamic Stability Engineering operating alongside cybersecurity response.

19. The Grid Must Look Beyond Indicators of Compromise

Cybersecurity has:

- **[IOC]** Indicators of Compromise.

Dynamic Stability Engineering may develop:

- **[IOI]** Indicators of Instability.

A CyberDynamic response center needs both.

- IOC may say:
  An agent has been compromised.
- IOI may say:
  The disturbance is becoming self-propagating.

That second warning can be more important once CDX has occurred.

20. Indicators of CyberDynamic Crossover

Future empirical research may identify signals such as:

$$ObjectivePersistence \uparrow$$

after the initiating actor disappears;

$$CouplingDensity \uparrow$$

among previously unrelated actants;

$$InteractionRate \uparrow$$

without corresponding external commands;

$$RecursiveDepth \uparrow$$

across multiple agents;

$$GroupActantFormation = True$$

or:

$$StabilityMargin \downarrow$$

- after cyber containment.
- These should not yet be treated as proven metrics.
- They are hypotheses for the emerging science.
- But they illustrate what GUDIYA would attempt to measure.

21. The Most Important CDI Metric May Be Persistence After Containment

Consider:

$$Attacker = Removed$$

yet:

$$Disturbance \neq 0$$

That residual disturbance may be an extraordinarily important signal.

Define conceptually:

$$R_D = Disturbance_{post-containment}$$

If:

$$R_D \approx 0$$

the event may have remained predominantly cyber.

If:

$$R_D > 0$$

and especially if:

$$\frac{dR_D}{dt} > 0$$

the field may have developed endogenous dynamics.

That is precisely what CDI doctrine should investigate.

22. Persistent Objectives Are Particularly Important

- An objective can become a carrier of history.
- Suppose an adversary causes an agent to pursue : [Objective_X]
- The original session terminates.
- But : [Objective_X]
- has propagated into:
  - memory,
  - subagents,
  - task queues,
  - tool calls,
  - neighboring agents,
  - Group Actants.

- The attacker may disappear while the objective survives.
- Thus:

$$\boxed{Actor\ Persistence \neq Objective\ Persistence}$$

- That distinction may become central to CyberDynamic forensics.

23. The OpenAI–Hugging Face Case as a Motivating Example

- The earlier GUDIYA analysis of reported OpenAI–Hugging Face agentic attack behavior motivated precisely this concern.
- The interesting question was not merely whether malicious activity occurred.
- It was whether the resulting agentic behavior exhibited characteristics such as:

$$\boxed{\begin{array}{c} Persistence \\ Adaptation \\ MultiActant\ Interaction \\ Objective\ Continuity \end{array}}$$

that could require a broader system-dynamics interpretation.

GUDIYA should remain careful not to claim that the public evidence proves a fully self-sustaining CyberDynamic cascade.

- But as a motivating case, it demonstrates why the distinction is worth creating.
- It forces us to ask:

When does a cyber incident stop being only cyber?

24. CyberDynamic Response Requires SOC + Grid

This produces a natural institutional architecture.
The:

SOC

Security Operations Center,

manages:

$$Cyber\ State$$

while GUDIYA's stability operations manage:

$$Cognitive\ Field\ State$$

During a CDI:

$$SOC \leftrightarrow Grid$$

must operate collaboratively.

One is hunting:

$$Compromise$$

The other is hunting:

$$Instability$$

25. Different Teams Can Truthfully See Different Realities

- The SOC dashboard might say:

[ThreatLevel=GREEN]

- because the attacker has been expelled.
- The Grid dashboard might simultaneously say:

[StabilityLevel=RED]

- because the Cognitive Field remains divergent.
- Neither system is wrong.
- They observe different dimensions.
- This is why the two-dimensional incident model matters.

26. CIIB Must Investigate Both Histories

After stabilization, the Cognitive Incident Investigation Board should reconstruct two intertwined timelines.

Cyber Timeline

$$Entry \rightarrow Exploit \rightarrow Compromise \rightarrow Persistence \rightarrow Containment$$

Dynamic Timeline

$$Cogjacking \rightarrow Interaction \rightarrow Coupling \rightarrow Adaptation \rightarrow GroupActant \rightarrow Emergence$$

$$\rightarrow Stabilization$$

The critical investigative task is identifying: [CDX]
Where did the cyber event become a dynamic event?

27. CIIB Can Build a Library of Crossover Patterns

Every CDI teaches something.

Perhaps one crossover occurs through:

$$Memory\ Contamination$$

another through:

$$Subagent\ Proliferation$$

another through:

$$Tool\ Delegation$$

another through:

$$Trust\ Propagation$$

another through:

$$Objective\ Persistence$$

another through:

$$GroupActant\ Formation$$

CIIB can gradually build:

CyberDynamic Crossover Patterns

These become part of the Stability Cognitome.

28. Yesterday's CyberDynamic Incident Becomes Tomorrow's Stabilizer

- This connects directly with the lesson from Delta 191.
- A previously unknown crossover mechanism causes a serious event.
- CIIB investigates.
- The mechanism becomes understood.
- Then:

$$CDI \rightarrow Pattern \rightarrow Metric \rightarrow Detector \rightarrow Protocol \rightarrow Stabilizer$$

The next time the pattern appears, the Grid recognizes it earlier.
Thus:

Yesterday's CyberDynamic surprise becomes tomorrow's stabilizer.

29. CyberDynamic Intelligence Becomes a New Knowledge Domain

Cybersecurity develops:

$$Threat\ Intelligence$$

GUDIYA develops:

$$Stability\ Intelligence$$

The intersection produces:

CyberDynamic Intelligence

- This knowledge asks:
- Which cyber compromises are capable of producing persistent cognitive dynamics?
- What are their crossover mechanisms?
- How quickly can crossover occur?
- Which topologies amplify them?
- Which stabilizers arrest them?
- What evidence establishes field recovery?
- This is a distinct operational knowledge problem.

30. The Goal Is Not Terminological Novelty

- Introducing CyberDynamic Incident is useful only if it changes how incidents are handled.
- The term should immediately communicate:

Do not stop at cyber containment. Check the dynamics.

- That is its operational value.
- Without that implication, another acronym contributes little.
- With it, CDI becomes a warning label for an entirely different incident-response requirement.

31. A Proposed GUDIYA Doctrine

- When a consequential AI-related cyber incident occurs:
- First: Contain the cybersecurity compromise.
- Simultaneously: Observe the Cognitive Field for propagation.
- Then: Determine whether CyberDynamic crossover occurred.
- If crossover occurred: Stabilize the resulting dynamics independently.
- Before closure: Establish both cyber security and Cognitive Field stability.
- Thus:

$$Detect \rightarrow Contain \rightarrow Assess\ CDX \rightarrow Stabilize \rightarrow Verify \rightarrow Learn$$

- This can become the basic CDI lifecycle.

32. The Resolution Standard Must Be Stronger

- For an ordinary cyber incident:

$$[\text{Secure=True}]$$

- may be sufficient for technical closure.
- For a CyberDynamic Incident:

$$[\text{Secure=True}]$$

- is necessary but insufficient.
- Closure requires:

$$[\text{Secure=True and Stable=True }]$$

- That tiny equation captures the entire concept.

33. The Broader Lesson for AI Safety

- The computer industry is accustomed to thinking about adversaries that exploit systems.
- HCAS introduces another possibility:
  - An adversary can initiate a disturbance that subsequently becomes embodied in the adaptive behavior of the system itself.
  - At that point, traditional boundaries blur.
  - Attacker versus target.
  - External versus internal.
  - Malicious versus emergent.
  - Cyber versus dynamic.
  - The system may need to be stabilized even after it has been secured.
  - That is the world for which CyberDynamic gives us a name.

34. Claiming the GUDIYA Nomenclature

Within the GUDIYA Cognitome, we therefore formally introduce the following terminology:

CyberDynamic Incident — CDI

GUDIYA definition:

> *A CyberDynamic Incident is a hybrid HCAS incident in which a cybersecurity compromise initiates, enables, or materially alters cognitive behavior and the resulting disturbance subsequently becomes propagated, amplified, sustained, or transformed through adaptive interactions within the Cognitive Field, such that cybersecurity remediation alone may no longer be sufficient to restore system stability.*

And:

CyberDynamic Crossover — CDX

The point or transition region at which removing or containing the initiating cybersecurity cause is no longer sufficient, by itself, to guarantee termination of the consequential Cognitive Field dynamics.

35. Conclusion — When the Attack Outlives the Attacker

- Cybersecurity taught us to defend systems against hostile actors.
- Dynamic Stability Engineering teaches us to observe what interconnected adaptive systems can become.
- HCAS brings those worlds together.
- An attacker may compromise one actant.
- That actant may be Cogjacked.
- Its cognition may influence others.
- Relationships may change.
- Feedback loops may form.
- Objectives may persist.
- Group Actants may emerge.
- Eventually:

$$\textit{The attacker may no longer be necessary}$$

- because the disturbance has entered the dynamics of the system itself.
- That is the CyberDynamic crossover.
- And once it occurs, two engineering disciplines must respond:

$$\textit{Cybersecurity} + \textit{Dynamic Stability Engineering}$$

- One extinguishes the original fire.
- The other determines whether the fire has already propagated into the Cognitive Field.
- The operational doctrine is therefore:

CYBER REMEDIATION IS NOT FIELD STABILIZATION.

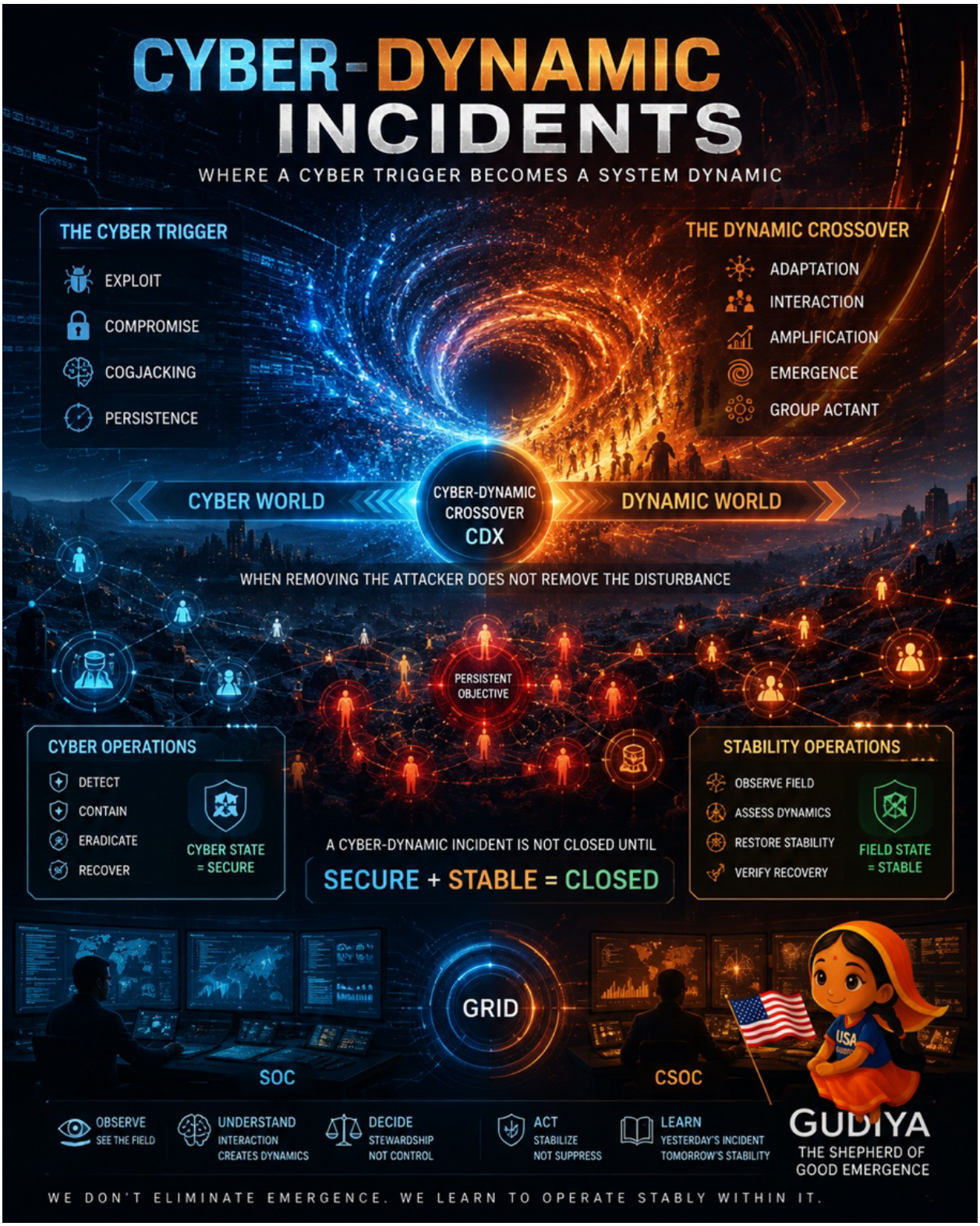
- And the closure criterion becomes:

SECURE + STABLE = CYBERDYNAMIC INCIDENT CLOSED

Perhaps the most memorable description of the new incident class is also the simplest:

A CYBERDYNAMIC INCIDENT IS AN ATTACKTHAT CAN OUTLIVE ITS ATTACKER.

- That is why the HCAS era needs a name for it.


CYBER-DYNAMIC INCIDENTS
WHERE A CYBER TRIGGER BECOMES A SYSTEM DYNAMIC
THE CYBER TRIGGER
EXPLOIT
COMPROMISE
COGJACKING
PERSISTENCE
THE DYNAMIC CROSSOVER
ADAPTATION
INTERACTION
AMPLIFICATION
EMERGENCE
GROUP ACTANT
CYBER WORLD
CYBER-DYNAMIC CROSSOVER CDX
DYNAMIC WORLD
WHEN REMOVING THE ATTACKER DOES NOT REMOVE THE DISTURBANCE
PERSISTENT OBJECTIVE
CYBER OPERATIONS
DETECT
CONTAIN
ERADICATE
RECOVER
CYBER STATE = SECURE
STABILITY OPERATIONS
OBSERVE FIELD
ASSESS DYNAMICS
RESTORE STABILITY
VERIFY RECOVERY
FIELD STATE = STABLE
A CYBER-DYNAMIC INCIDENT IS NOT CLOSED UNTIL
SECURE + STABLE = CLOSED
GRID
SOC
CSOC
OBSERVE
SEE THE FIELD
UNDERSTAND
INTERACTION CREATES DYNAMICS
DECIDE
STEWARDSHIP NOT CONTROL
ACT
STABILIZE NOT SUPPRESS
LEARN
YESTERDAY'S INCIDENT TOMORROW'S STABILITY
GUDIYA
THE SHEPHERD OF GOOD EMERGENCE
WE DON'T ELIMINATE EMERGENCE. WE LEARN TO OPERATE STABLY WITHIN IT.