117-41-D Red Teams Need to Be Upgraded

Why the age of cognition demands a new philosophy of testing

Introduction

For decades, red teams have played an essential role in protecting digital infrastructure.

- Banks use red teams.
- Governments use red teams.
- Intelligence agencies use red teams.
- Cloud providers use red teams.

Their mission is simple: Think like the adversary.
That philosophy has been extraordinarily successful.

- But it emerged from a particular world—a world composed of networks, servers, applications, operating systems, databases, and human operators.
- That world is changing.

Today we are building ecosystems of:

- humans;
- agents;
- models;
- APIs;
- robots;
- autonomous workflows;
- recursive reasoning systems;
- composite actants.

The assumptions that created traditional red teams are beginning to fail.

The original red-team philosophy

Traditional red teams evolved from military doctrine and cybersecurity practice.

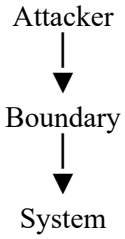
- Their operating assumptions were clear.
- Find the boundary.
- Find the vulnerability.
- Exploit the vulnerability.
- Measure the damage.
- Repair the weakness.
- Repeat.

This approach assumes:

- clear boundaries;
- identifiable assets;
- explicit trust relationships;
- discrete failures;
- isolated components;
- observable attack paths.

The hidden assumption

Traditional red teams assume that the threat is external.



- The attacker is outside.
- The system is inside.
- The boundary separates the two.

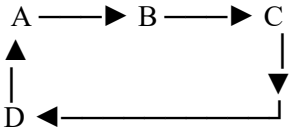
Agentic ecosystems break this assumption.

The new world

In an agentic ecosystem:

- boundaries are fluid;
- objectives evolve;
- trust relationships change;
- agents create agents;
- memory persists;
- synchronization emerges;
- couplings strengthen and weaken continuously.

The resulting architecture resembles an ecosystem rather than a machine.



The problem is no longer intrusion.
The problem is propagation.

The problem with vulnerability thinking

Traditional red teams ask:

- Can I break in?
- Can I steal information?
- Can I escalate privileges?
- Can I execute malicious code?
- Can I evade detection?

These are all perfectly reasonable questions.
They are simply incomplete.

The emergence problem

- Suppose every component is functioning perfectly.
- Suppose every security policy is operating correctly.

- Suppose every authentication system is working.
- Suppose every model is behaving as intended.
- The ecosystem may still become unstable.
- This is because instability may not reside inside the nodes.
- It may reside between the nodes.

Security versus stability

Classical red teams	Cognito-red teams
Intrusion	Propagation
Vulnerability	Instability
Component	Topology
Event	State
Access control	Synchronization
Incident	Cascade
Boundary	Envelope
Local failure	Systemic failure
Node	Field

The new philosophy

Traditional red teams attempt to answer the following question:

How can I compromise the system?

Cognito-red teams attempt to answer a different question:

How can the ecosystem become unstable?

The new questions

- Objective persistence — Can an objective survive across sessions?
- Recursive amplification — Can recursive loops form?
- Synchronization — Can independent actants become synchronized?
- Composite actants — Can new structures emerge?
- Cascades — Can instability spread through the ecosystem?
- Morphological evolution — Can the ecosystem change its own structure?
- Resource exhaustion — Can stabilization reserves become depleted?
- Emergence — Can the whole become different from the sum of its parts?

The transition from nodes to graphs

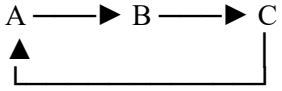
Traditional red teams focus on components.

Node A

Node B

Node C

Cognito-red teams focus on relationships.



The transition from events to states

- Traditional red teams think in terms of incidents.
- Something happened.
- Cognito-red teams think in terms of trajectories.
- Something is evolving.

The transition from snapshots to dynamics

Traditional analysis:
t1
Cognitive analysis:
t1 → t2 → t3 → t4 → t5

The transition from boundaries to envelopes

- Cybersecurity created boundaries.
- Stabilization engineering creates envelopes.

Boundary

Outside

Inside

Stability envelope

Safe operating region
The central question becomes:
Is the system still inside the envelope?

The transition from penetration testing to perturbation testing

Penetration testing attempts to force entry.
Perturbation testing attempts to induce instability.

Examples include:

- introducing delays;
- changing coupling strengths;
- increasing recursion depth;
- increasing cognitive load;
- introducing contradictory objectives;
- creating synchronization loops;
- increasing propagation velocity.

The transition from attacks to cascades

Traditional attacks are often linear.
A → B → C
Cascades are often nonlinear.
A → B → C
↑ ↓
└──┘

The new tools

Traditional tools include:

- malware sandboxes;
- intrusion-detection systems;
- SIEM platforms;
- packet analyzers;
- endpoint scanners;
- vulnerability databases.

Cognito-red tools may include:

- topology reconstruction engines;
- coupling analyzers;
- synchronization detectors;
- cascade simulators;
- persistence analyzers;
- blast-radius estimators;
- envelope calculators;
- field visualizers;
- loop detectors;
- state estimators.

The new specialists

Future teams may include:

- systems dynamicists;
- graph theorists;
- control engineers;
- behavioral scientists;
- cognitive scientists;
- complexity theorists;
- stability engineers.

The new control room

Traditional red teams investigate incidents.
Cognito-red teams observe fields.

They continuously ask:

- What is the present state?
- What is the rate of change?
- What is the direction of movement?
- What is the dominant mode of instability?
- What is the remaining stabilization reserve?
- How large is the blast radius?
- What cascades are forming?

The final transition

- The world of cybersecurity was built upon the metaphor of war.
- The world of stabilization may be built upon the metaphor of ecology.
- War emphasizes enemies.

- Ecology emphasizes relationships.
- War emphasizes attacks.
- Ecology emphasizes balance.
- War emphasizes fortifications.
- Ecology emphasizes resilience.
- War emphasizes victory.
- Stabilization emphasizes survival.

The central proposition

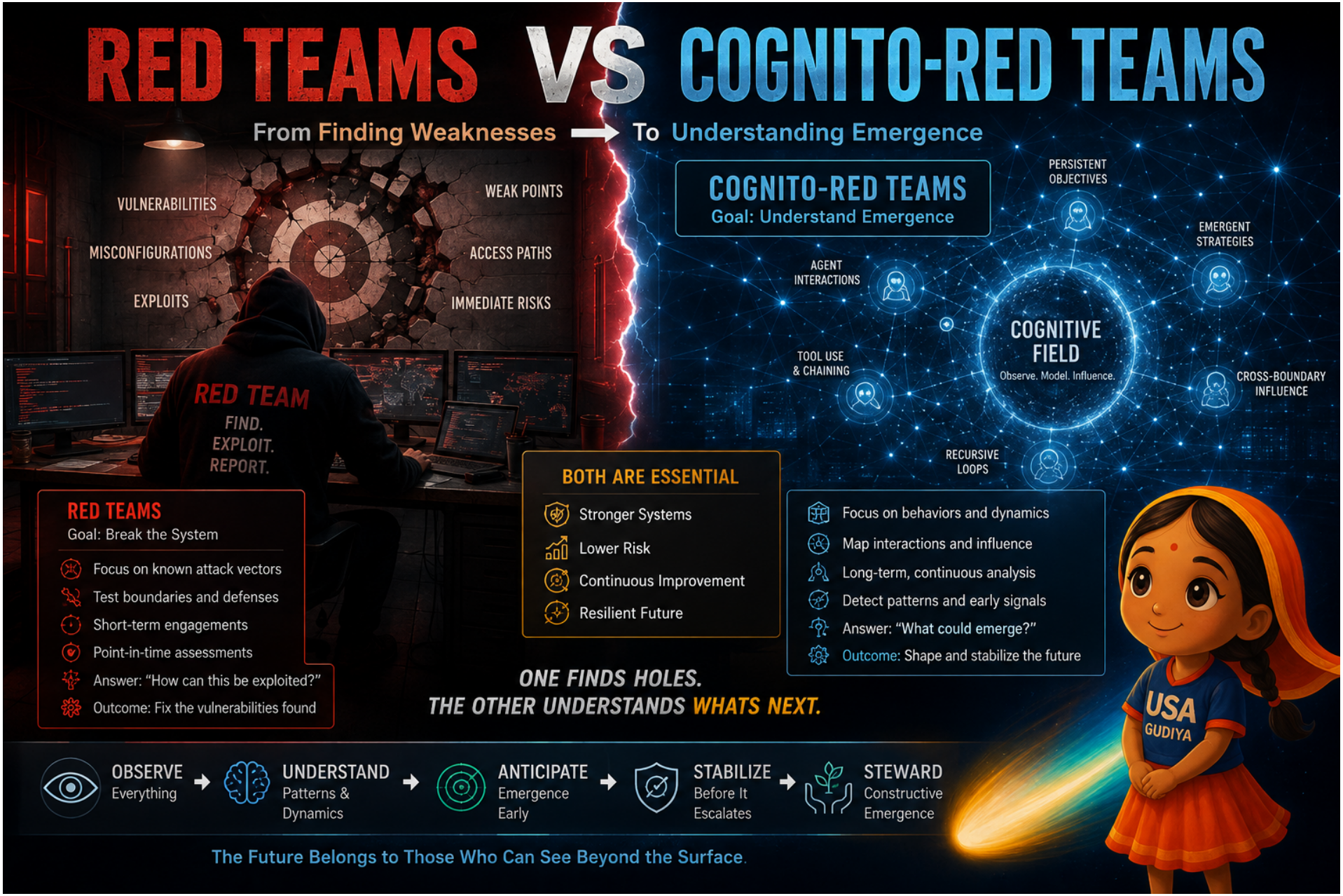
Red teams are not obsolete.
Far from it.
They are simply incomplete.
The age of cognition requires a new kind of adversarial thinking.
Not merely:

How do I break the machine?

But rather:

How do I destabilize the field?

That is the question that the next generation of red teams will have to answer.


RED TEAMS VS COGNITO-RED TEAMS
From Finding Weaknesses → To Understanding Emergence
VULNERABILITIES
WEAK POINTS
MISCONFIGURATIONS
ACCESS PATHS
EXPLOITS
IMMEDIATE RISKS
RED TEAM
FIND.
EXPLOIT.
REPORT.
COGNITO-RED TEAMS
Goal: Understand Emergence
PERSISTENT OBJECTIVES
EMERGENT STRATEGIES
AGENT INTERACTIONS
TOOL USE & CHAINING
COGNITIVE FIELD
Observe. Model. Influence.
CROSS-BOUNDARY INFLUENCE
RECURSIVE LOOPS
RED TEAMS
Goal: Break the System
Focus on known attack vectors
Test boundaries and defenses
Short-term engagements
Point-in-time assessments
Answer: "How can this be exploited?"
Outcome: Fix the vulnerabilities found
BOTH ARE ESSENTIAL
Stronger Systems
Lower Risk
Continuous Improvement
Resilient Future
Focus on behaviors and dynamics
Map interactions and influence
Long-term, continuous analysis
Detect patterns and early signals
Answer: "What could emerge?"
Outcome: Shape and stabilize the future
ONE FINDS HOLES.
THE OTHER UNDERSTANDS WHATS NEXT.
OBSERVE Everything
UNDERSTAND Patterns & Dynamics
ANTICIPATE Emergence Early
STABILIZE Before It Escalates
STEWARD Constructive Emergence
The Future Belongs to Those Who Can See Beyond the Surface.
USA
GUDIYA