TACTICAL NEW ORDER
MANAGED CYBER PROTECTION

TACTICAL

# YELLOW TEAM

AI SECURITY DELIVERED AS A PRACTICE

CYBER SECURITY THAT STOPS HACKERS

WWW.NEWORDERGROUP.NET

ISO/IEC 27001 & ISO9001

// 01 // AI SECURITY PRACTICE

# THE EXPOSURE YOU HAVE ALREADY CREATED

You did not just add a chatbot. You wired AI into how the business runs. It approves transactions, updates records, moves data, and ships code. It holds credentials, calls your systems, and acts on its own at machine speed with no human watching each step.

## THE AI THAT ANSWERS IS A POLICY PROBLEM. THE AI THAT ACTS IS A SECURITY PROBLEM.

Policy, training, and data loss controls govern what a person types into an assistant. They cannot see the agent your team wired into a workflow last month, they cannot test how it behaves when an attacker manipulates it, and they cannot stop it acting in real time. Every agent is a non human identity you never onboarded and never offboarded. Classic application testing assumes predictable software. Agents are non deterministic, and they open entirely new attack classes.

// THE NUMBERS BEHIND THE EXPOSURE

| METRIC | FIGURE | SOURCE |
|---|---|---|
| Enterprise applications expected to run task specific AI agents by end 2026 | **40%, up from under 5% a year earlier** | GARTNER 2025 |
| Organisations that have already adopted agents | **79%** | ENTERPRISE AI ADOPTION SURVEYS 2025 |
| Organisations with no AI governance policy | **63%** | ENTERPRISE AI ADOPTION SURVEYS 2025 |
| AI breaches hitting organisations with no AI access controls | **97%** | IBM COST OF A DATA BREACH 2025 |
| Breaches now featuring shadow AI | **20%** | IBM COST OF A DATA BREACH 2025 |
| Added to the average breach cost where shadow AI is present | **USD 670,000** | IBM COST OF A DATA BREACH 2025 |

// NEW ATTACK CLASSES AGENTS OPEN

### PROMPT INJECTION

Hostile instructions smuggled into the content an agent reads.

### AGENT HIJACKING

An attacker redirects the agent to act on their behalf.

### TOOL ABUSE

The agent's own system access turned into the weapon.

### DATA EXFILTRATION

Sensitive data drawn out through the agent's reach.

// 02 // AI SECURITY PRACTICE

# WHAT YELLOW TEAM IS

Cyber security has always split along colour lines. Red attacks. Blue defends. Purple makes the two work together. Yellow is the side that builds and runs the systems. AI moved your real exposure onto the build side, so we built a practice to match it.

| TEAM | FUNCTION | SCOPE |
|---|---|---|
| RED | OFFENCE | Attacks your estate to prove what an adversary can reach and what it would cost you. |
| BLUE | DEFENCE | Monitors, detects, and responds to what is already moving inside the environment. |
| PURPLE | INTEGRATION | Closes the loop between offence and defence, turning proven attacks into detections that hold. |
| **YELLOW** | **BUILD AND RUN** | **Secures what your business builds and operates. Today that means models, agents, tools, prompts, and the pipelines that ship them.** |

NEWORDER Yellow Team is our AI security practice. One team is accountable for discovering, testing, protecting, and governing every AI agent, model, and tool in your environment. It is human led, continuously run, and technology backed.

> **YELLOW TEAM IS NOT A SCAN AND NOT A REPORT. IT IS AN ACCOUNTABLE PRACTICE WITH A NAME, A SCOPE, AND A REPORTING LINE.**

// THE FOUR PILLARS

## 01

### DISCOVER

Inventory every AI agent, model, tool, prompt, and guardrail across cloud, SaaS, and CI/CD, including the shadow agents nobody registered.

// CONTINUOUS THREAT EXPOSURE MANAGEMENT

## 02

### TEST

Red team each agent before and after deployment against real adversary techniques, mapped to the OWASP LLM and Agentic Top 10.

// ADVERSARY PATH ENGINEERING

## 03

### PROTECT

Block prompt injection and agent hijacking at runtime, with human operators validating what the tooling flags.

// HUMAN VALIDATION AT SCALE

## 04

### GOVERN

Map posture to recognised standards and report it in language your board, your auditor, and your regulator can act on.

// EXECUTIVE CYBER RISK MANAGEMENT

// 03 // AI SECURITY PRACTICE

# WHAT WE DELIVER

Six modules, delivered as one managed service. Take the whole practice or start with the module that closes your largest gap.

// THE MODULES

// MODULE 01

## AI DISCOVERY AND AI BOM

A live bill of materials for your AI. Every agent, model, tool, system prompt, and guardrail mapped across cloud platforms, SaaS, and CI/CD pipelines, including homegrown applications.

// MODULE 02

## AI SECURITY POSTURE MANAGEMENT

Misconfiguration and policy gap analysis, supply chain risk assessment, and exposure scoring aligned to NIST AI RMF and OWASP.

// MODULE 03

## AGENTIC RED TEAMING

Multi turn adversarial testing against a live attack library. Context poisoning, tool chain manipulation, agent hijacking, and prompt injection, run before and after deployment.

// MODULE 04

## RUNTIME PROTECTION

Policy enforcement at the proxy, API, or AI gateway layer. Hostile instructions are blocked in the path of the agent, before it acts, not reported afterwards.

// MODULE 05

## AI DETECTION AND RESPONSE

Behavioural anomaly monitoring and intent analysis, with NEWORDER operators triaging every alert so your team acts on what is real.

// MODULE 06

## GOVERNANCE AND BOARD REPORTING

Standards mapped posture reporting, incident disclosure readiness, and executive briefings delivered through Executive Cyber Risk Management.

// STANDARDS ALIGNMENT

## MAPPED. TESTED. REPORTABLE.

NIST AI RMF · OWASP TOP 10 FOR LLM · MITRE ATLAS · ISO/IEC 27001 · EU AI ACT

NIST AI RMF for governance, the OWASP Top 10 for LLM applications for testing, MITRE ATLAS for adversary technique mapping, ISO 27001 for the management system, and EU AI Act readiness for regulatory reporting.

// 04 // AI SECURITY PRACTICE

# THE LASSO SECURITY PARTNERSHIP

NEWORDER Yellow Team is delivered on the Lasso Security platform. Lasso builds AI security for the enterprise: discovery, posture management, automated red teaming, runtime enforcement, and detection and response, purpose built for agents rather than retrofitted from application security tooling.

| 300K+ | 100% | 500+ | 50+ |
|---|---|---|---|
| ATTACK PAYLOADS IN THE RED TEAM LIBRARY | OWASP LLM AND AGENTIC TOP 10 COVERAGE | NEW ATTACK VARIANTS ADDED WEEKLY | EVASION TECHNIQUES EXERCISED |

## // WHY WE SELECTED LASSO

| | |
|---|---|
| BUILT FOR AGENTIC AI | The platform was designed around agents that act, not chat interfaces that answer. |
| ENFORCEMENT IN THE PATH | Controls run at the proxy, API, and AI gateway layer, so a hostile instruction is stopped before the agent acts on it. |
| ADVERSARIAL TESTING AT SCALE | A payload library of more than 300,000 attacks, multi turn, covering context poisoning and tool chain manipulation, with full MITRE and OWASP LLM and Agentic Top 10 coverage. |
| PRODUCTION GRADE PERFORMANCE | Sub 50 millisecond classification means the controls run in live systems without slowing the business down. |
| ENTERPRISE INTEGRATION | Native coverage across major cloud AI platforms and CI/CD pipelines, so shadow agents surface rather than hide. |

## // DIVISION OF RESPONSIBILITY

| LASSO PROVIDES | NEWORDER PROVIDES |
|---|---|
| The platform, engines, and attack library | **The operators, the scope, and the rules of engagement** |
| Discovery and AI BOM across your estate | **Adversary path engineering against your agents** |
| Runtime enforcement at the gateway layer | **Tuning, triage, and human validation of findings** |
| Detection, intent analysis, and telemetry | **Board reporting and the accountability for outcomes** |

// THE SPLIT THE PLATFORM FINDS IT. OUR OPERATORS DECIDE WHAT IT MEANS AND WHAT YOU DO ABOUT IT.

// 05 // AI SECURITY PRACTICE

# LEAVE BEHIND TECHNOLOGY LASSO

Most assessments leave you a report and a follow up quote. Yellow Team leaves you a running control. When the engagement ends, the Lasso platform stays in your environment, deployed, tuned, and enforcing, with your team trained to operate it.

// WHAT STAYS, AND WHAT IT MEANS FOR YOU

**01**
**LIVE AI INVENTORY**
A continuously updated AI bill of materials covering every agent, model, tool, and guardrail. It keeps discovering after we leave, so next quarter's shadow agent surfaces on its own.

**02**
**RUNTIME ENFORCEMENT**
Policies live at your gateway and proxy layer, blocking prompt injection and hijack attempts in real time, tuned to your traffic rather than shipped as defaults.

**03**
**CONTINUOUS RED TEAMING**
The attack library keeps running against your agents on a schedule you control. Testing becomes a standing control instead of an annual event.

**04**
**DETECTION AND RESPONSE**
Behavioural monitoring and intent analysis wired into your existing security operations, with alerting your analysts already know how to consume.

**05**
**AUDIT READY EVIDENCE**
Standards mapped reporting you can hand to an auditor, a regulator, an enterprise client, or the board without rebuilding it by hand.

| <50MS | <200MS | <5SEC | 1.4% |
|---|---|---|---|
| REAL TIME BLOCKING AT THE GATEWAY | TIME TO DETECT AN AI THREAT | FULL SESSION ANALYSIS | FALSE POSITIVE RATE |

// TRANSFER OF CONTROL

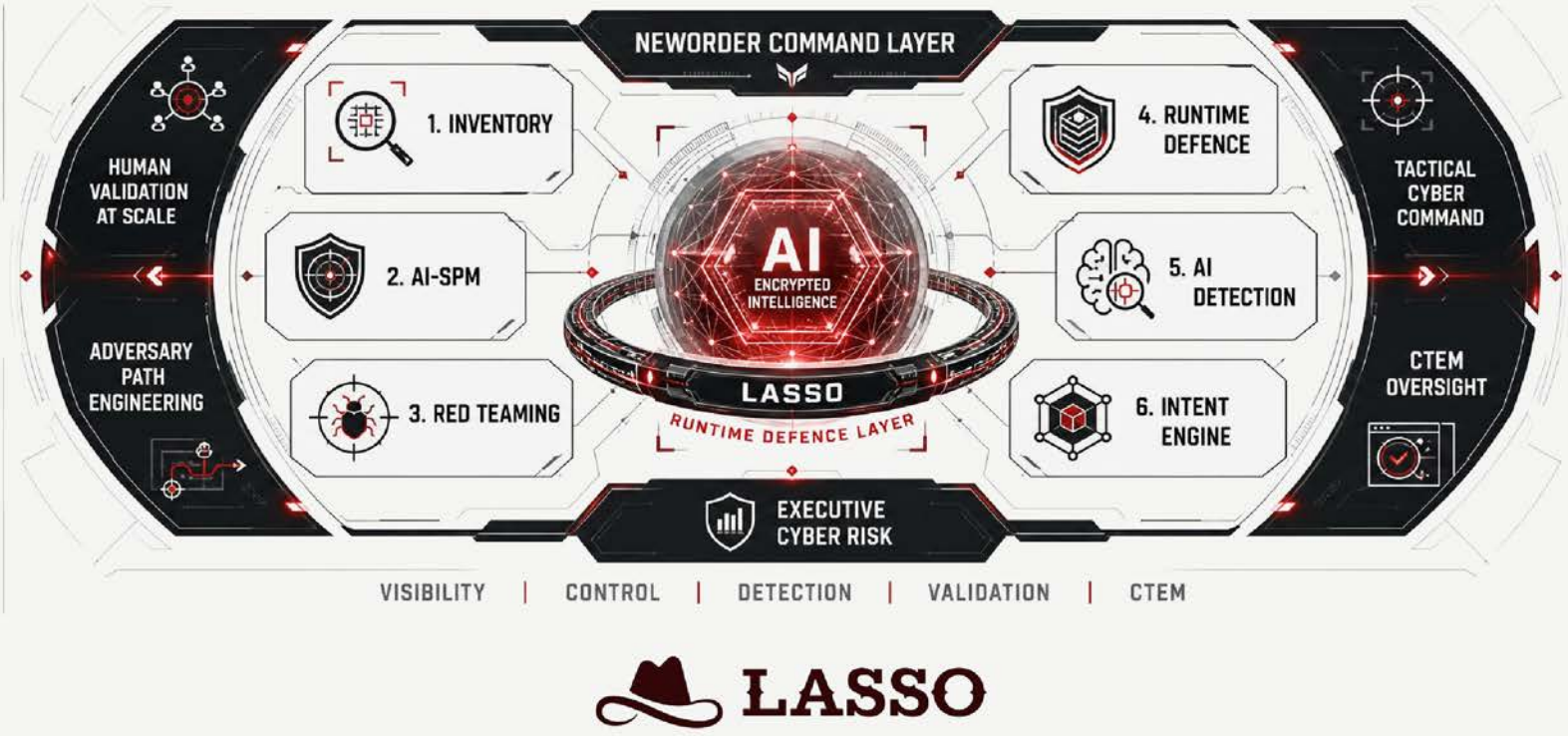
## YOU DO NOT RENT THE OUTCOME. YOU KEEP THE CONTROL.

The technology is deployed in your environment, tuned to your traffic, transferred to your team, and supported by NEWORDER.

// 06 // AI SECURITY PRACTICE

# HOW IT FITS TOGETHER

Six modules operating inside one command structure. NEWORDER runs the command layer and the human validation. The Lasso platform carries runtime defence in the path of every agent.




LASSO

// PLATFORM COVERAGE

AWSBedrock · OpenAI · CopilotStudio · Anthropic
GoogleVertex AI · ServiceNow · AzureAI Foundry · Microsoft 365 Copilot
Salesforce · Salesforce Agentforce · Power Platform · and more

// THE STACK VISIBILITY, CONTROL, DETECTION, VALIDATION, AND CTEM OVERSIGHT IN ONE PRACTICE.

// 07 // AI SECURITY PRACTICE

# HOW THE ENGAGEMENT RUNS

Five phases. The first is a thirty minute conversation. The last one never ends, because your AI estate does not stand still.

## // THE FIVE PHASES

| | | | |
|---|---|---|---|
| 01 | DISCUSSION | 30 MINUTES | A no-obligation discussion to understand your requirement and to provide guidance on how NEWORDER can assist. No deployment required. |
| 02 | DEPLOY | WEEKS 1 TO 2 | The platform is stood up in your environment. Discovery runs across cloud, SaaS, and CI/CD to produce your first AI bill of materials. |
| 03 | TEST | WEEKS 3 TO 4 | Adversarial testing against every agent in scope. Findings are ranked by business consequence, not by scanner severity. |
| 04 | PROTECT | WEEK 4 ONWARD | Runtime policies are enforced at the gateway. Detection is wired into your security operations and tuned against live traffic. |
| 05 | RUN | CONTINUOUS | Managed by NEWORDER operators. Reported monthly. Reviewed quarterly with your executive team and your board. |

## // WHAT YOU RECEIVE

### AI BILL OF MATERIALS

Every agent, model, tool, and guardrail in your estate, kept current.

### HIJACK PATH REPORT

Ranked adversary routes to your agents, with the fix for each one.

### BOARD PACK

Posture, exposure, and progress in language the board already uses.

// DELIVERY MANAGED BY NEWORDER OPERATORS. REPORTED MONTHLY. REVIEWED QUARTERLY WITH YOUR BOARD.

// 08 // AI SECURITY PRACTICE

# WHY ENGAGE NEWORDER

You can buy a platform. You can hire a tester. You can appoint an integrator. Yellow Team is all three, under one accountable team, with one number to call.

// SIX REASONS

## 01

### ONE ACCOUNTABLE TEAM

You do not assemble a tool, an integrator, and a tester, then manage the gaps between them. Yellow Team owns the outcome end to end.

## 02

### OFFENSIVE HERITAGE

NEWORDER is an offensive security business. We engineer the adversary's path to your agents, then close it. Testing is our starting point, not an add on.

## 03

### HUMAN VALIDATION AT SCALE

Every finding is triaged by an operator before it reaches you. Your team spends its time on what is real instead of on alert volume.

## 04

### TECHNOLOGY YOU KEEP

The Lasso platform stays behind, deployed and tuned, not withdrawn at the end of the engagement. The control outlives the project.

## 05

### BOARD READY GOVERNANCE

Posture mapped to NIST AI RMF, OWASP, and MITRE ATLAS, reported through Executive Cyber Risk Management in language your board acts on.

## 06

### REGULATORY COVER

Evidence for POPIA, GDPR, the EU AI Act, and the rapid incident disclosure duties that now hold executives answerable for the gap.

// THE POSITION

## OUTCOME OVER OUTPUT.

Accountability for agentic AI has already moved to the executive. The exposure is live, the controls are absent in most organisations, and the disclosure clock starts the moment something goes wrong. Yellow Team closes that gap and leaves the control running.

// 09 // AI SECURITY PRACTICE

# WHO YELLOW TEAM IS FOR

Yellow Team is built for organisations that have already put AI to work inside the business, and now carry the consequence of it.

## // SECTORS

### FINANCIAL SERVICES

Agents touching payments, onboarding, and client records under regulators who expect rapid disclosure.

### INSURANCE

Agents pricing risk and handling claims, where a manipulated decision becomes a liability of its own.

### HEALTHCARE AND MEDICAL

Agents handling special personal information, where a hijacked workflow is notifiable on its own.

### RETAIL, ECOMMERCE AND LOGISTICS

Agents moving orders, payments, stock, and supplier data at machine speed across systems nobody inventories.

### TECHNOLOGY AND SOFTWARE

Coding agents with repository and pipeline access, where tool abuse becomes your customers' supply chain problem.

### ONLINE GAMING AND GAMBLING

Agents inside player accounts, payouts, and fraud controls, where abuse converts straight into cash.

### MARITIME

Agents in vessel operations, chartering, and port logistics, where an instruction moves physical cargo.

### MINING

Agents across operational technology, plant reporting, and supply chain, where downtime is the loss event.

### TELECOMMUNICATIONS

Agents with subscriber data and network provisioning access, held to national critical infrastructure duties.

## // ENGAGE YELLOW TEAM WHEN

- Your teams have deployed AI agents and nobody holds the full list. //
- An agent in your environment holds credentials to a system that matters. //
- Your board has asked what your AI exposure is and the answer took more than a day. //
- An enterprise client, insurer, or regulator has started asking for AI assurance evidence. //
- You are about to ship an AI product and want it tested before your customers test it for you. //

// THE NEXT STEP

## A THIRTY MINUTE DISCUSSION.

A no-obligation discussion to understand your requirement and to provide guidance on how NEWORDER can assist. No deployment, no commitment.

BOOK DISCUSSION →

EMAIL INFO@NEWORDERGROUP.NET · WEB WWW.NEWORDERGROUP.NET