JODOR Technologies

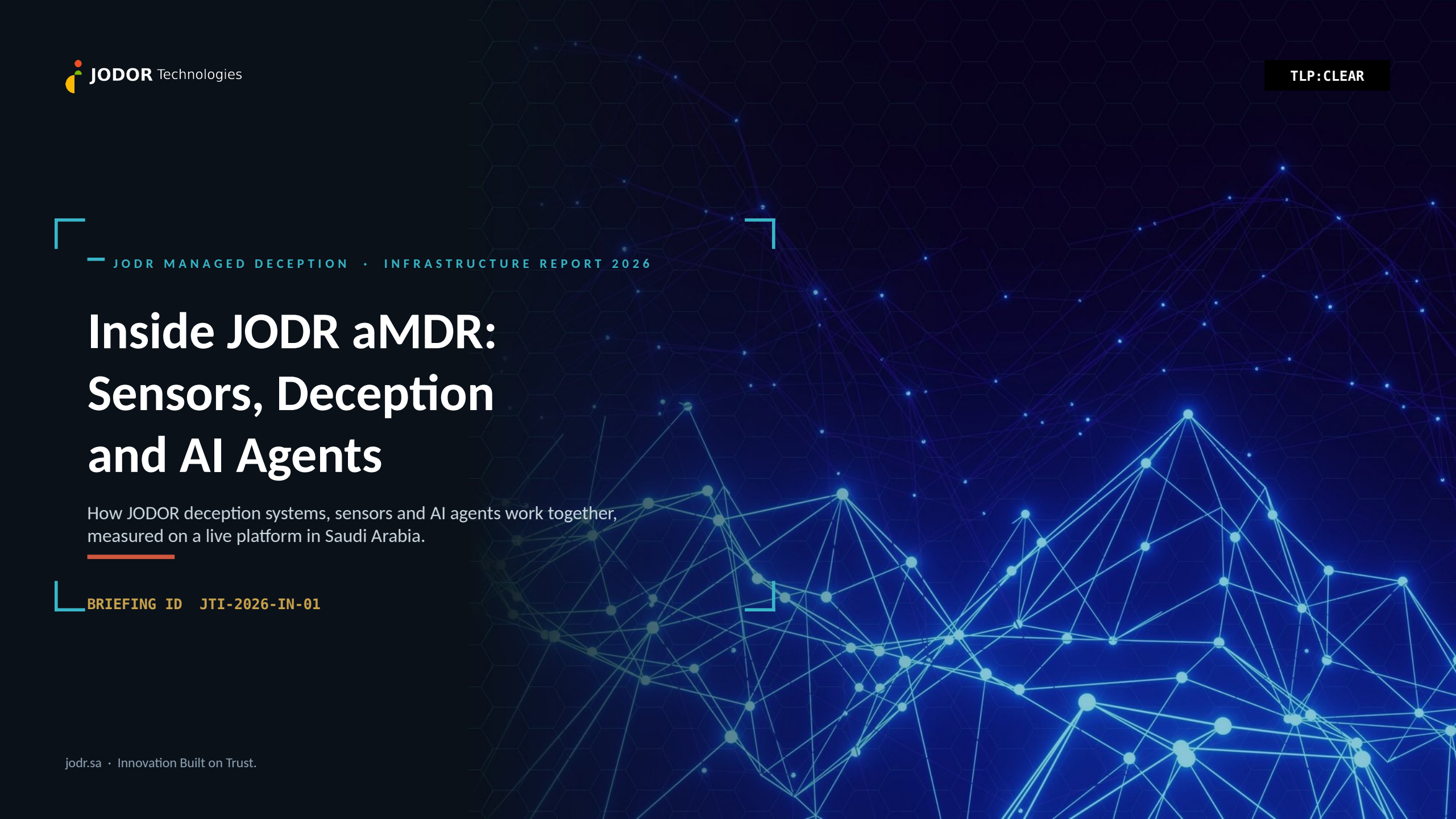
TLP:CLEAR

JODR MANAGED DECEPTION · INFRASTRUCTURE REPORT 2026

Inside JODR aMDR: Sensors, Deception and AI Agents

How JODOR deception systems, sensors and AI agents work together, measured on a live platform in Saudi Arabia.

BRIEFING ID JTI-2026-IN-01



INFRASTRUCTURE REPORT

Key findings

Four agents shared the work, and each alert went only as far as it needed to.

The evidence: Over 15 days, JODOR ran AI agents with a human in the loop on a live platform in Saudi Arabia, and 6,168 alerts were raised.

KEY FINDINGS

82% — **The AI agents settled most alerts alone.**
They settled 5,073 of the 6,168 alerts without passing them to a person.

41% — **Two in five alerts needed no AI model.**
The rule agent settled 2,506 alerts on proof that already existed.

16% — **A person saw only the alerts that needed one.**
The agents escalated 967 alerts to the human in the loop.

64% — **Most blocks rested on proof, not judgement.**
The rule agent decided 1,234 of the 1,920 blocks.

5% — **Speed still needs safety checks.**
Of 1,906 blocks applied, 102 were later undone; safety checks caught 87.

WHAT TO DO

1. **Let rules block attacks that are already proven.**
2. **Send unclear alerts to an expert AI model.**
3. **List your trusted addresses before you automate.**

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

THE SETUP

Attacks now outrun manual defence

JODOR sensors and industry research show how little time defenders now have.

84%

Of 173 attackers who used a known exploit, 146 used it within 10 seconds of arriving.

27 s

The fastest attacker in one public study spread through a network 27 seconds after breaking in.

28%

A typical security operations centre (SOC) leaves 28% of its alerts uninvestigated, a 2026 survey found.

JODOR assessment: people alone almost certainly cannot respond within 10 seconds, so defence now needs machines as well.

Sources: JODOR sensors and deception systems in Saudi Arabia (JODOR analysis; raw data not published) · CrowdStrike 2026 Global Threat Report, press release (24 Feb 2026) · Prophet Security, State of AI in the SOC 2026 (3 Aug 2026)

THE SETUP

Where this data comes from

Context and definitions used throughout this report.

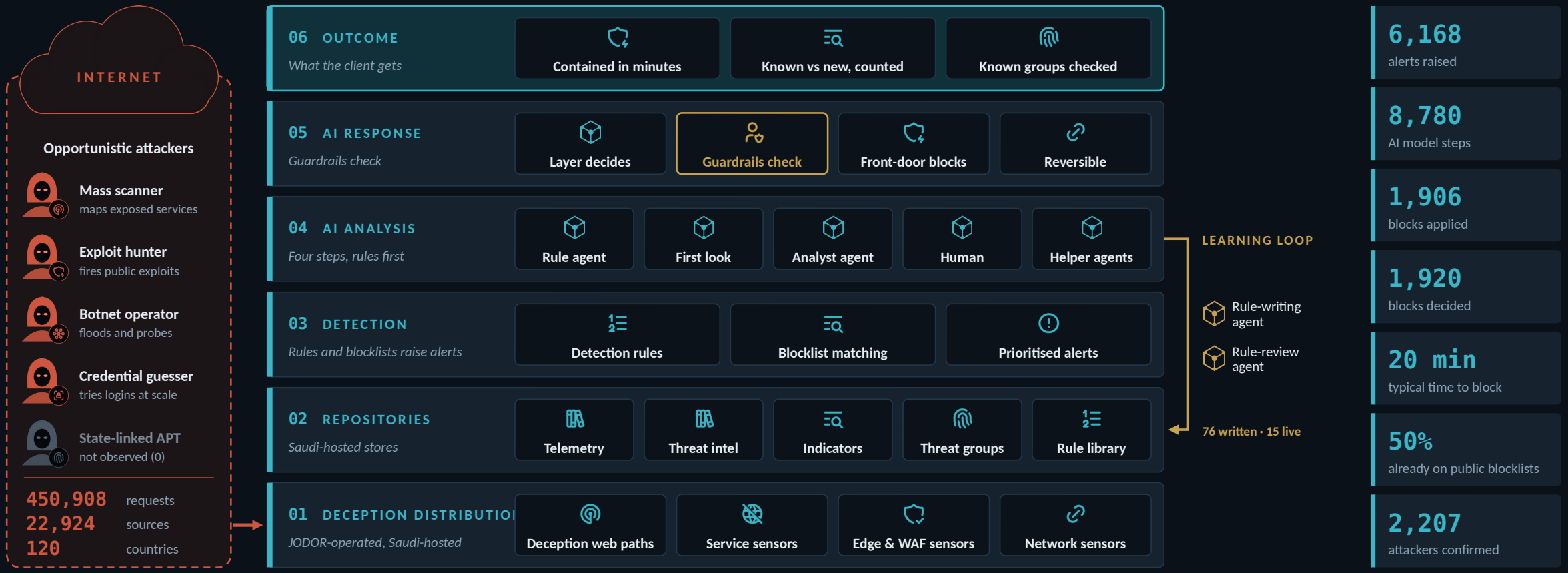
Term	Definition	Count
Sensors	Four types (deception web paths, service sensors, edge and web application firewall sensors, network sensors) in front of one live JODOR-operated web platform in Saudi Arabia, not a bank	56 days
Request	One web request logged with address, time, path and browser string	450,908
Visitor address	A distinct IPv4 or IPv6 source address	22,924
Attacker	An address JODOR confirmed hostile, or one that sent a known-exploit or secret-file request	2,207
Known-exploit attacker	Sent a request matching a published exploit pattern	173
Secret hunter	Asked for secret or configuration files (.env, .git/config, cloud keys, SSH keys)	261
On a blocklist	Matched at least one of 34 public blocklists	1,105
Country	First country recorded for the address (hosting location, not attribution)	84

Sources: JODOR sensors and deception systems in Saudi Arabia (JODOR analysis; raw data not published)

THE SETUP

Sensor and engine architecture

Read from the bottom: attackers hit the deception systems, and each stage turns raw traffic into a decision.



6,168 alerts raised

8,780 AI model steps

1,906 blocks applied

1,920 blocks decided

20 min typical time to block

50% already on public blocklists

2,207 attackers confirmed

Sources: JODOR sensors and deception systems in Saudi Arabia (JODOR analysis; raw data not published) · JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

THE SETUP

How the observation ran

Sensors captured traffic for 56 days, and the AI agents were measured over 15 of those days.

TIMELINE (DURATIONS ONLY)

Sensor capture · 56 days

AI agents measured · 15 days

4

Sensor types deployed, built to look vulnerable, in front of one live JODOR-operated web platform in Saudi Arabia.

JODOR assessment: the agents acted alone, inside safety checks, for all 15 days. Every alert and block is counted once.

Sources: JODOR sensors and deception systems in Saudi Arabia (JODOR analysis; raw data not published) · JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

HOW THE AGENTS WORK

Agents act; people set policy

Agentic AI means AI agents that act on what they find, not only advise.



AI agents

They read each alert, judge it, and block the attacker when the evidence is clear.



Safety checks

Fixed checks stop the agents from blocking administrators, partners, delivery networks and real search engines.



Human in the loop

A security professional sets the policy and handles every alert the agents escalate.

JODOR assessment: judge any AI defence by what its agents may do alone and what they must pass to a person.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

HOW THE AGENTS WORK

AI built the defence system

AI engineering agents wrote the software and its rules, following the policy people set.

163

AI agents wrote 163 detection rules, each one for a defined type of attack.

7

Seven blocking rules act only when a security control or blocklist has already proven the attack.

15

A rule-writing agent has since added 15 detection rules, each checked first by an AI reviewer.

JODOR assessment: people decided what counts as proof, and AI turned each decision into a rule that acts in seconds.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

HOW THE AGENTS WORK

Four steps, one alert

An alert moves up a step only when the step below cannot settle it.

STEP 4

Human in the loop

A security professional handles the alerts the agents escalate.

Human

STEP 3

Analyst agent

A large expert AI model (117 billion parameters) investigates what the first look could not settle.

18 s

STEP 2

First-look agent

A fast AI model reads the alert and settles the clear cases.

1.2 s

STEP 1

Rule agent

It blocks only when a security control or blocklist has already proven the attack. It needs no AI model.

No model

HELPER AGENTS

A fix agent applies recommended fixes after an AI reviewer approves them.

A rule-writing agent proposes new detection rules for review.

A profiling agent groups alerts by attacker.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

WHAT HAPPENED

Where each alert was settled

Each of the 6,168 alerts is counted once, at the step that settled it.



Step 1 · Rule agent	2,506	41%
Step 2 · First-look agent	1,305	21%
Step 3 · Analyst agent	1,262	20%
Step 4 · Human in the loop	967	16%
Still open when the 15 days ended	128	2%

JODOR assessment: known threats never reached an AI model, so the models were kept for the alerts that needed judgement.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

WHAT HAPPENED

The rule agent needs proof

The rule agent settles an alert only when one of these four conditions is true.

1. The source is already blocked, or a security control has already denied it.
2. The firewall has denied every request the source sent.
3. The source is on a public blocklist and was denied, or it sent 10 or more requests.
4. The same alert was judged before, or it comes from a known administrator.

If none of the four conditions is true, the alert moves up to the first-look agent.

JODOR assessment: AI wrote these rules, but each one acts on evidence a control has already proven, never on a guess.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

WHAT HAPPENED

Rule agent block conditions

Each block by the rule agent met one fixed condition, and no AI model was asked at that moment.

Condition	Blocks
Log sweep: a firewall-denied attack signature; or a blocklist listing with a deny or 10 or more requests; or a Tor exit with a deny	537
Indicator store: earlier real alerts, firewall signatures, blocklists or indicators named by an AI analysis	288
Firewall denied every request from the source	248
Range merge: 3 or more blocked attackers in one /24 network within 7 days	115
Edge rules: a probed path or scanner user agent never served to real users	22
Repeat attacker: 3 or more real alerts in 24 hours	21
Hosting network: 10 or more attacking addresses and no legitimate requests	3
Total blocks by the rule agent	1,234

JODOR assessment: 309 blocks (indicator store 288, repeat attacker 21) reused earlier AI findings. No minimum AI confidence was set; safety checks were the control.

Sources: JODOR response data (JODOR analysis)

WHAT HAPPENED

Expert AI for unclear alerts

Alerts without proof go to a fast AI model first, then to an expert model if still unclear.

1.2 s

The first-look agent typically reads and settles an alert in 1.2 seconds.

18 s

The analyst agent typically completes a deep investigation in 18 seconds.

JODOR assessment: the fast model handles the volume, so the expert model is kept for the alerts that need depth.

Sources: JODOR AI analysis and rules data (JODOR analysis)

WHAT HAPPENED

One in six alerts was escalated

Alerts the agents could not settle were escalated to the human in the loop.

967

The agents escalated 967 of the 6,168 alerts (16%) to the human in the loop.

5,073

The agents settled the other 5,073 alerts (82%) without a person.

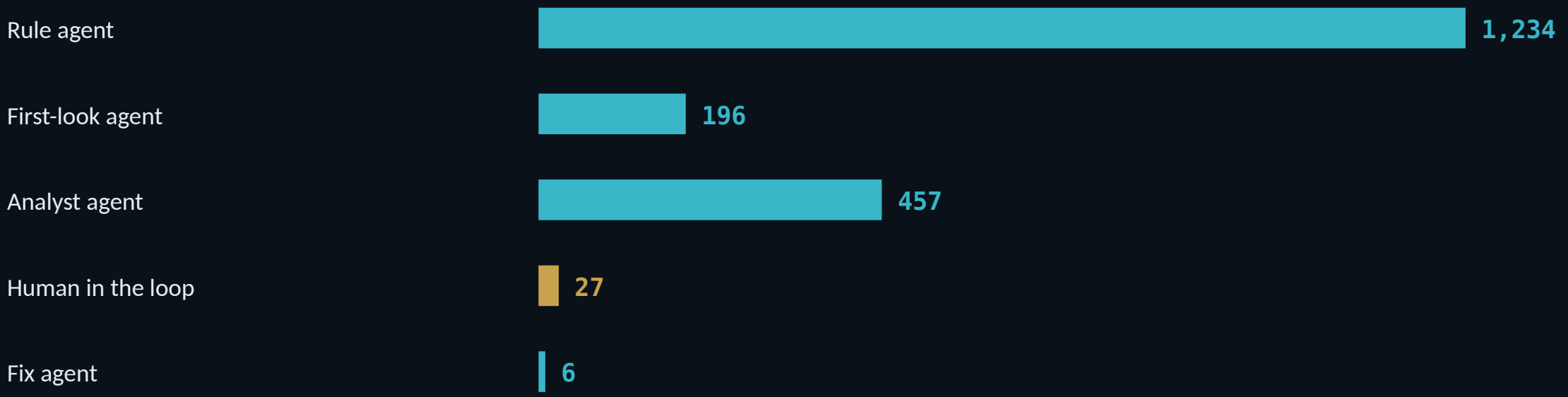
JODOR assessment: the last 128 alerts (2%) were still open when the 15 days ended. A person saw only what needed a person.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR response data (JODOR analysis)

WHAT HAPPENED

Most blocks rested on proof

Each of the 1,920 blocks is counted once, under the agent that decided it.



JODOR assessment: 309 of the rule agent's blocks reused earlier AI findings stored as known attacker addresses. In the end, 1,906 blocks were applied.

Sources: JODOR response data (JODOR analysis)

WHAT HAPPENED

The firewall first, the agents next

How long it took from an attacker’s first request to a permanent block.

5 min

When the firewall had already denied every request, the rule agent blocked within about 5 minutes.

14 min

When the analyst agent decided, the block typically followed within 14 minutes.

20 min

Across all 1,759 blocks that could be timed, the typical time was 20 minutes.

15 s

The fastest block came 15 seconds after the attacker’s first request.

JODOR assessment: the firewall stopped the first request, and the agents then closed the door for good. The first request may have been harmless, so real times are shorter.

Sources: JODOR response data (JODOR analysis)

WHAT HAPPENED

Three analysts' work became three hours

What 6,168 alerts in 15 days would mean in people's time.

3

Working every alert by hand would need about 3 analysts every day.

64

The agents escalated about 64 alerts a day to the human in the loop.

3 h

Handling those escalations took about 3 hours of the human in the loop's day.

JODOR assessment: this assumes about 3 minutes of hands-on work per alert and 8-hour days. The freed time goes to threat hunting.

Sources: JODOR AI analysis and rules data (JODOR analysis) · JODOR calculation at about 3 minutes of hands-on work per alert, 8-hour days

WHAT HAPPENED

Fast is not always right

Some blocks were later undone. This page shows how many, and what undid them.

1,906

Of the 1,920 blocks decided, 1,906 were applied; 14 failed or were withdrawn.

102

Later, 102 of the applied blocks, about 5%, were undone.

87

Safety checks undid 87 of those blocks automatically.

15

The other 15 were undone after escalation to the human in the loop.

JODOR assessment: most wrong blocks hit administrators, delivery networks and real search engines. A list of trusted addresses would very likely prevent them.

Sources: JODOR response data (JODOR analysis)

LESSONS

Give freedom in step with proof

Ask one question about each alert: is the attack proven, and can the block be undone?

ACT ALONE

The rule agent may block alone when the attack is proven and the block can be undone.

ASK FIRST

The first-look and analyst agents propose risky blocks, and the human in the loop approves them.

HUMAN DECIDES

The human in the loop decides on partners, shared addresses and every escalated alert.

JODOR assessment: a second pair of eyes checks every risky block. In the test, AI agents also blocked alone, so start new AI decisions in the middle column.

Sources: JODOR response data (JODOR analysis) · Mohsin et al., human–AI collaboration in SOCs with trusted autonomy (2025)

HOW TO START

Start with one step

Two ways to begin, and how JODOR supports you after that.

01

JODOR deception

Deception in your name catches your attackers on first contact.

02

AI agents and an analyst

Rule, first-look and analyst agents, with a person in the loop, turn every touch into intelligence.

03

Professional services

JODOR experts help you harden systems, clean up secrets and act on the list.

START WITH ONE STEP

45 min A threat briefing on what now attacks organisations like yours.

30 days A free pilot with deception in your name that ends in a board-ready report.

Not ready yet? Forward this report to the head of your security operations.

Sources: JODOR sensors and deception systems in Saudi Arabia (JODOR analysis; raw data not published) · JODOR AI analysis and rules data (JODOR analysis)

JODOR Technologies

TLP:CLEAR

JODR MANAGED DECEPTION · INFRASTRUCTURE REPORT 2026

From managed detection to managed deception

WEB	jodr.sa
EMAIL	info@jodr.sa
LINKEDIN	linkedin.com/company/jodr-technologies
OFFICE	Riyadh, Kingdom of Saudi Arabia

Book a 45-minute briefing or a free 30-day pilot at info@jodr.sa.

Scan to email info@jodr.sa