AI IN THE SECURITY OPERATIONS CENTRE

PRACTITIONER INSIGHTS ON TRUST, AUTONOMY AND THE FUTURE OF DEFENSIVE DECISION-MAKING

WWW.RIVERSAFE.CO.UK

RIVERSAFE

CONTENTS

1. Executive summary 3

2. Before adoption: two critical perspectives 4

3. Where ai is delivering real operational advantage 6

4. Where leaders should apply restraint 10

5. The trust gap 11

6. The workforce shift 12

7. Governing intelligence without slowing innovation 13

8. Final thoughts 14

1. EXECUTIVE SUMMARY

Our work across complex enterprise environments provides a close view of how security operations are evolving in practice. Not in vendor demonstrations, but in live operational settings where the cost of being wrong is measured in disruption, regulatory exposure, and reputational damage.

What we are observing is not a wholesale transformation of the Security Operations Centre, but a deliberate recalibration. Organisations are not rushing blindly toward autonomy, instead they are experimenting, learning, and importantly drawing boundaries.

The question is no longer whether AI belongs in the SOC, the more consequential question is:

Where does AI reliably enhance human judgment and where would trusting it introduce unacceptable risk?

Organisations are reporting measurable value where work is volumetric, repetitive, and context-heavy. Yet hesitation remains strongest around areas requiring interpretation, accountability, and high-impact decision-making.

The future SOC will not be defined by the volume of intelligence it deploys, but by the quality of the decisions it enables. Speed alone is insufficient as, without judgment, acceleration introduces instability rather than resilience.

This is why we believe the defining leadership challenge is no longer visibility, it is decision velocity, the ability to interpret signals quickly, act with confidence, and do so within well-governed boundaries.

At the same time, a strategic workforce concern is emerging. As AI reduces Tier 1 and Tier 2 workload, leaders must confront an uncomfortable question: if junior analysts no longer develop investigative instincts, where will the next generation of senior practitioners come from?

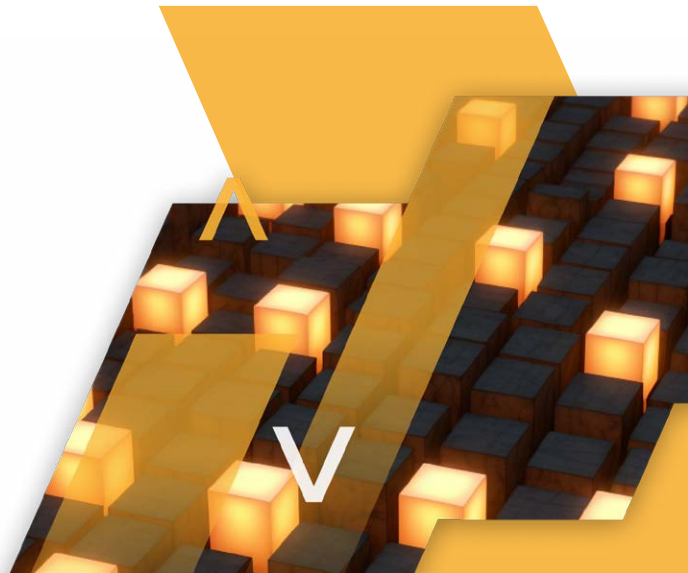
The organisations that succeed will not be those that automate the fastest but those that balance acceleration with discipline.

This paper reflects both RiverSafe's practitioner perspective working across multiple SOC environments and insights gathered from a roundtable discussion with heads of security operations and cyber leaders from across sectors. Under Chatham House rules, they shared their unvarnished experiences with AI-augmented security operations.

WHO THIS PAPER IS FOR

This paper is written for security leaders responsible for protecting complex organisations in an era of accelerating technological change.

It is not intended as a technical manual, nor as a commentary on emerging tooling. Rather, it offers a leadership perspective on how artificial intelligence is reshaping the operating model of cyber defence and what this means for organisations seeking to remain resilient.



2. BEFORE ADOPTION: TWO CRITICAL PERSPECTIVES

Before examining where AI is delivering operational value and where caution remains necessary, two foundational considerations shape how organisations should approach AI adoption in security operations.

1. THE SHIFT FROM VISIBILITY TO DECISION VELOCITY

For over a decade, cyber defence centred on closing visibility gaps. Today most organisations can see far more than they once could and the constraint has shifted with the modern SOC operating under decision pressure.

Analysts face overwhelming volumes of alerts, correlations, and competing interpretations.

What differentiates effective security operations is not awareness, but the speed and confidence with which teams interpret information and act upon it.

Artificial intelligence is fundamentally altering this dynamic by reducing the distance between detection and comprehension. But acceleration without judgement introduces risk.

RiverSafe refers to this construct as the Decision Velocity Model, a practical way for security leaders to calibrate autonomy against operational risk.

DECISION VELOCITY ZONES

Decision Velocity Zone	Primary Decision Maker	Typical Actions	Risk Profile	Leadership Posture
Assisted Velocity	Human-led	Alert enrichment, summaries	Low	Encourage adoption
Augmented Velocity	Human-in-the-loop	Triage recommendations	Moderate	Scale deliberately
Bounded Autonomy	AI within guardrails	False-positive closure	Elevated	Govern tightly
Full Autonomy	AI-led	Strategic response	High	Not yet advisable

2. AI IS NOT A SHORTCUT TO SECURITY MATURITY

A dangerous misconception is emerging in boardrooms that AI can compensate for operational weaknesses. However, AI is not a maturity engine, it is a maturity amplifier.

As one roundtable participant observed: “if your processes are messy or non-existent, AI has nothing to accelerate.”

Process maturity is a prerequisite for AI effectiveness. The fundamentals have not changed, just execution speed has.

The organisations seeing real value from AI share common characteristics. They have invested in foundational capabilities that AI can leverage. Without these in place, AI just amplifies the gaps.

Foundational Capability	Why It Matters
Identity discipline	Decisions rely on identity context
Asset intelligence	Unknown assets create risk
Detection hygiene	Weak detections scale noise
Segmentation	Limits blast radius
Incident readiness	Enables human intervention

3. WHERE AI IS DELIVERING REAL OPERATIONAL ADVANTAGE

Across RiverSafe engagements and industry observation, one pattern is consistently visible: AI delivers the greatest value when applied to problems characterised by scale, repetition, and cognitive load. In these environments, intelligence does not replace analysts, it enhances their ability to reach clarity faster and removes operational friction.

Operational advantage is increasingly defined not by who sees first, but by who understands first.

Following are the areas where AI is proving to add measurable improvements in real SOCs.

REDUCING TIME-TO-UNDERSTAND

One of the most immediate benefits of AI lies in compressing investigative timelines. Traditionally, analysts were required to pivot manually across endpoint telemetry, identity stores, network logs, threat intelligence feeds, and historical case data before forming a view of risk. This process was effective but slow.

AI is changing that dynamic by aggregating and contextualising signals automatically.

Instead of asking analysts to assemble the puzzle, intelligence presents a near-complete picture highlighting behavioural anomalies, mapping activity chains, and surfacing relevant historical patterns.

The result is a measurable reduction in what RiverSafe refers to as time-to-understand, an emerging metric that may soon rival mean-time-to-detect in strategic importance.

When analysts reach clarity faster, several downstream benefits follow:

- **Faster containment decisions**
- **Reduced attacker dwell time**
- **Lower cognitive fatigue**
- **Greater investigative confidence**

CONTEXT AGGREGATION

Modern SOCs are not short of telemetry, they are overwhelmed by it. AI is proving particularly effective at transforming fragmented data into coherent investigative context.

By correlating identity behaviour, asset criticality, privilege levels, geolocation anomalies, and prior alerts, AI allows analysts to move immediately toward risk evaluation rather than basic information gathering.

Security teams are transitioning from environments that are merely data-rich to ones that are genuinely insight-rich.

In practice, this means analysts can answer critical questions earlier in the investigation lifecycle:

- **Is this behaviour consistent with the user's profile?**
- **Does the affected asset carry elevated business risk?**
- **Has similar activity been observed previously?**
- **Does this align with known adversarial techniques?**

ANALYST AUGMENTATION

A significant portion of SOC workload has historically consisted of necessary but low-leverage tasks: documentation, case summarisation, evidence formatting, and internal reporting. While operationally important, these activities diverted attention from deeper analytical work.

AI is now absorbing much of this overhead. Automated case narratives, incident summaries, and executive-ready reporting are allowing analysts to redirect their expertise toward interpretation and decision-making the areas where human judgment is most valuable.

DETECTION ENGINEERING LIFT

Detection engineering has always been one of the highest-leverage activities within cyber defence, yet it has traditionally been constrained by time and specialist skill availability.

AI is beginning to change this equation. Emerging capabilities can analyse historical alert patterns, identify detection gaps, recommend logic refinements, and assist in translating threat intelligence into actionable detection rules.

One practitioner shared an example of using AI to write queries. They had built a private integration with Google Gemini, ingesting their organisation's field names and indexes so analysts could generate SIEM queries using natural language.

The productivity gains were immediate - queries that once took twenty minutes or longer to construct now appeared in seconds. However, they had deliberately kept the AI outside their SIEM environment, requiring human review before execution. Their caution proved warranted when Gemini returned a query that was almost perfect, except it had misspelled 'dedupe' with an extra letter, which would have caused the query to fail.

Importantly, AI does not reduce the importance of detection engineers, it increases it. AI accelerates the learning loop, but human expertise remains essential for validating defensive logic and understanding adversary tradecraft.

The organisations deriving the most value are those treating AI as a collaborator rather than a replacement. Weak detections amplified by AI simply generate noise faster whereas strong detections scale protection.

IMPROVING SIGNAL-TO-DECISION RATIOS

Alert fatigue remains one of the most persistent challenges in security operations. High volumes of low-confidence alerts dilute analyst focus and increase the risk that genuinely critical signals are overlooked.

AI is demonstrating growing capability in filtering benign patterns, clustering related alerts, and identifying which signals merit immediate attention. The strategic outcome is not fewer alerts but better prioritisation.

This improves what might be described as the signal-to-decision ratio: the proportion of alerts that genuinely require human judgment. As this ratio improves, so too does the effectiveness of the SOC.

VULNERABILITY INTELLIGENCE

Another emerging advantage lies in vulnerability analysis. Many organisations struggle not with identifying vulnerabilities, but with prioritising them meaningfully.

AI can assist by correlating vulnerability data with exploit intelligence, asset exposure, business criticality, and environmental context shifting conversations from raw severity scores toward operational risk narratives.

This enables leaders to answer a far more strategic question: Which vulnerabilities actually matter right now? When remediation decisions become risk-informed rather than volume-driven, defensive posture strengthens.

TECHNICAL OPPORTUNITIES ON THE NEAR HORIZON

While today's gains are already meaningful, several technical opportunities are rapidly approaching operational viability.

PREDICTIVE THREAT MODELLING

AI-driven behavioural analysis may soon allow organisations to anticipate likely attack paths based on environmental patterns moving defence incrementally from reactive toward anticipatory.

ADAPTIVE PLAYBOOKS

Static response playbooks could evolve into adaptive workflows that adjust dynamically based on incident context, asset sensitivity, and threat actor behaviour.

AUTONOMOUS DATA TRIAGE

As telemetry volumes continue to expand, AI may increasingly determine which data streams warrant deeper inspection, reducing storage, processing, and investigative overhead.

SECURITY KNOWLEDGE GRAPHS

The fusion of identity, asset, behavioural, and threat intelligence into unified knowledge structures represents one of the most promising frontiers for contextual defence.

These capabilities remain emergent, yet their trajectory is clear.

ASPIRATIONAL OPPORTUNITIES

Looking further ahead, AI has the potential to reshape the operating philosophy of cyber defence.

Rather than reacting to alerts, mature organisations may operate within continuously evaluated risk states environments where intelligence constantly reassesses exposure and recommends preventative adjustments.

Security operations could become less event-driven and more posture-driven. Decision-making may evolve from episodic investigation toward continuous risk calibration.

However, aspiration must remain grounded in realism. Autonomy should expand only as governance, auditability, and organisational trust mature in parallel.

4. WHERE LEADERS SHOULD APPLY RESTRAINT

Technological capability does not automatically translate into operational readiness. While artificial intelligence is advancing at remarkable speed, experienced security leaders understand that maturity is defined as much by what an organisation chooses not to automate as by what it does.

Several domains within cyber defence still demand careful human oversight not because machines lack sophistication, but because the consequences of error remain disproportionately high.

HIGH-IMPACT RESPONSE ACTIONS

High-impact response actions sit firmly within this category. There was clear consensus among roundtable participants: nobody in the room was prepared to let AI make high-impact decisions like shutting down environments.

Fully autonomous containment such as disabling identity platforms, isolating critical infrastructure, or triggering enterprise-wide controls, introduces a level of operational risk that most organisations are not yet structured to absorb. Even low-probability errors become unacceptable when the scale of disruption is large.

THREAT INTERPRETATION AND INTELLIGENCE

AI-generated outputs are often highly plausible, which can create a false sense of assurance. In cybersecurity, however, near-accuracy is insufficient. Investigations frequently hinge on nuance - contextual signals that separate anomalous from malicious behaviour.

As practitioners noted, threat interpretation, adversary intent analysis, and root cause evaluation remain deeply human disciplines.

AI can describe what has occurred with increasing precision, but understanding why it occurred and what it implies for future exposure continues to rely on human experience.

DATA SOVEREIGNTY AND CONTROL

Data sovereignty presents another boundary where caution is warranted. As one security leader put it: “the biggest threat from AI adoption is the risk of sensitive data leaving their environment.”

Security telemetry represents one of the most sensitive concentrations of organisational data. Leaders are rightly scrutinising where that data is processed, who can access it, and what residual exposure may persist.

5. THE TRUST GAP

Security professionals are trained sceptics. Their effectiveness depends on questioning assumptions, validating signals, and maintaining a posture of constructive doubt. AI introduces systems that are probabilistic by design capable of producing valuable insight, yet not always able to articulate their reasoning in ways that satisfy traditional expectations of evidence.

AI asks security teams to accept outputs they cannot fully verify, from systems they cannot fully explain, making recommendations they cannot always trace back to first principles.

Within the SOC, trust is not granted lightly, the cost of misplaced confidence is measured in operational disruption, regulatory scrutiny, and reputational harm. Defensive decisions must be explainable, defensible, and auditable. AI therefore faces a higher bar than most technologies, it must be reliable under pressure, not just impressive in demos.

Trust is built gradually through consistent performance and transparent behaviour. Practitioners were not hostile to AI, they were already using it and finding value. But they drew a sharp line between comfort and trust. They are comfortable experimenting, comfortable letting AI enrich alerts or draft reports. They are not comfortable letting AI make decisions that matter.

Leaders should view this trust gap not as an obstacle, but as a natural stage in technological maturity. Managing it requires clarity of communication, visible governance, and importantly, the preservation of human accountability.

Over time, as intelligence proves itself, the boundary between human-led and machine-assisted decision-making will shift. But that shift should be earned through evidence, not driven by enthusiasm.

6. THE WORKFORCE SHIFT

Much of the AI conversation has centred on tooling and automation. Far less attention has been given to what may ultimately prove the more consequential impact: the reshaping of the security workforce.

Practitioners at the roundtable were clear about workforce changes already underway. AI is reducing routine Tier 1 work although 'career enrichment rather than replacement' captured the intent. However, Tier 1 headcount has declined, making entry into the profession more competitive.

Higher tiers are seeing workload reduction from AI and MSSPs, with work shifting toward oversight, validation, and handling cases AI cannot resolve.

Four concerns emerged from the roundtable:

- **Reduced exposure.** If AI handles routine tickets, juniors lose the repetitive practice that builds foundational skills. The tedious BAU work is often where analysts develop pattern recognition and intuition.
- **AI-native graduates.** Participants questioned whether graduates who relied heavily on AI tools at university truly understand fundamentals, not just how to do something, but why certain activities are risky.
- **The pipeline problem.** As one practitioner put it, organisations risk 'creating a long-term problem for themselves', future senior practitioners may lack the depth of experience needed to identify issues and make informed decisions.
- **Judgment cannot be taught by AI.** The contextual understanding that separates good analysts from great ones comes from experience, not from reviewing AI outputs.

Forward-looking organisations are already adapting, preserving manual investigative opportunities within training environments and redesigning career pathways to emphasise analytical reasoning, threat modelling, and contextual interpretation earlier.

The skills profile of the future SOC is likely to evolve accordingly. Curiosity, critical thinking, and the ability to challenge machine-generated conclusions may become more valuable than procedural familiarity with tools.

Efficiency gains should therefore be balanced against long-term capability building. The goal is not simply a leaner SOC, but a smarter and more resilient one.

7. GOVERNING INTELLIGENCE WITHOUT SLOWING INNOVATION

As AI becomes embedded in security workflows, governance must evolve from procedural oversight to architectural intent.

The challenge is not choosing between control and innovation, but designing an environment where both reinforce one another.

Accountability

Effective governance begins with clarity of accountability. Machines may recommend, correlate, and even execute within defined boundaries, but responsibility for defensive outcomes must remain unequivocally human. This principle is foundational not only for operational integrity but also for regulatory defensibility.

AI systems should increasingly be treated as critical infrastructure subject to the same expectations of resilience, monitoring, and access control as other high-impact platforms. Their influence over decision-making warrants nothing less.

Autonomy boundaries

Not every action requires human approval, yet the parameters within which AI operates should be transparent, auditable, and aligned to organisational risk tolerance. AI systems should be treated as critical infrastructure, subject to the same expectations of resilience, monitoring, and access control as other high-impact platforms.

Data Stewardship

Minimising unnecessary exposure, understanding processing pathways, and maintaining visibility into how sensitive telemetry is handled are essential safeguards in an AI-augmented environment.

Auditability

Leaders should be able to answer, with confidence, why a recommendation was made, how a decision was reached, and what controls governed that outcome. This must be designed in, not retrofitted.

When approached deliberately, governance enables intelligence to scale safely, without slowing innovation.

8. FINAL THOUGHTS

AI is already reshaping how organisations detect threats, conduct investigations, and make defensive decisions. Yet amid this acceleration, constants remain. Fundamentals still matter, governance is not optional, and human judgment remains irreplaceable.

The organisations that derive lasting advantage will not be those that automate indiscriminately, nor those that hesitate excessively. They will be the ones that adopt AI deliberately, accelerating where it strengthens defensive posture, applying restraint where trust has yet to be earned, and maintaining clear accountability throughout.

The defining advantage is no longer technological capability. It is the confidence with which decisions are made under pressure. Because while AI is increasing the tempo of cyber defence, disciplined execution determines outcomes. And in an environment where attackers are learning quickly, disciplined organisations will remain the hardest to compromise.

Thank you to the security leaders who shared their insights at the RiverSafe AI in SOC Roundtable.

READY TO EMPOWER YOUR TEAMS AND ACCELERATE YOUR SOFTWARE DELIVERY?
Let RiverSafe help you build a platform that drives innovation, reduces complexity, and delivers measurable value.
Get in touch with us today to learn how we can help transform your developer experience.
www.riversafe.co.uk
Enquiries@RiverSafe.co.uk
RIVERSAFE