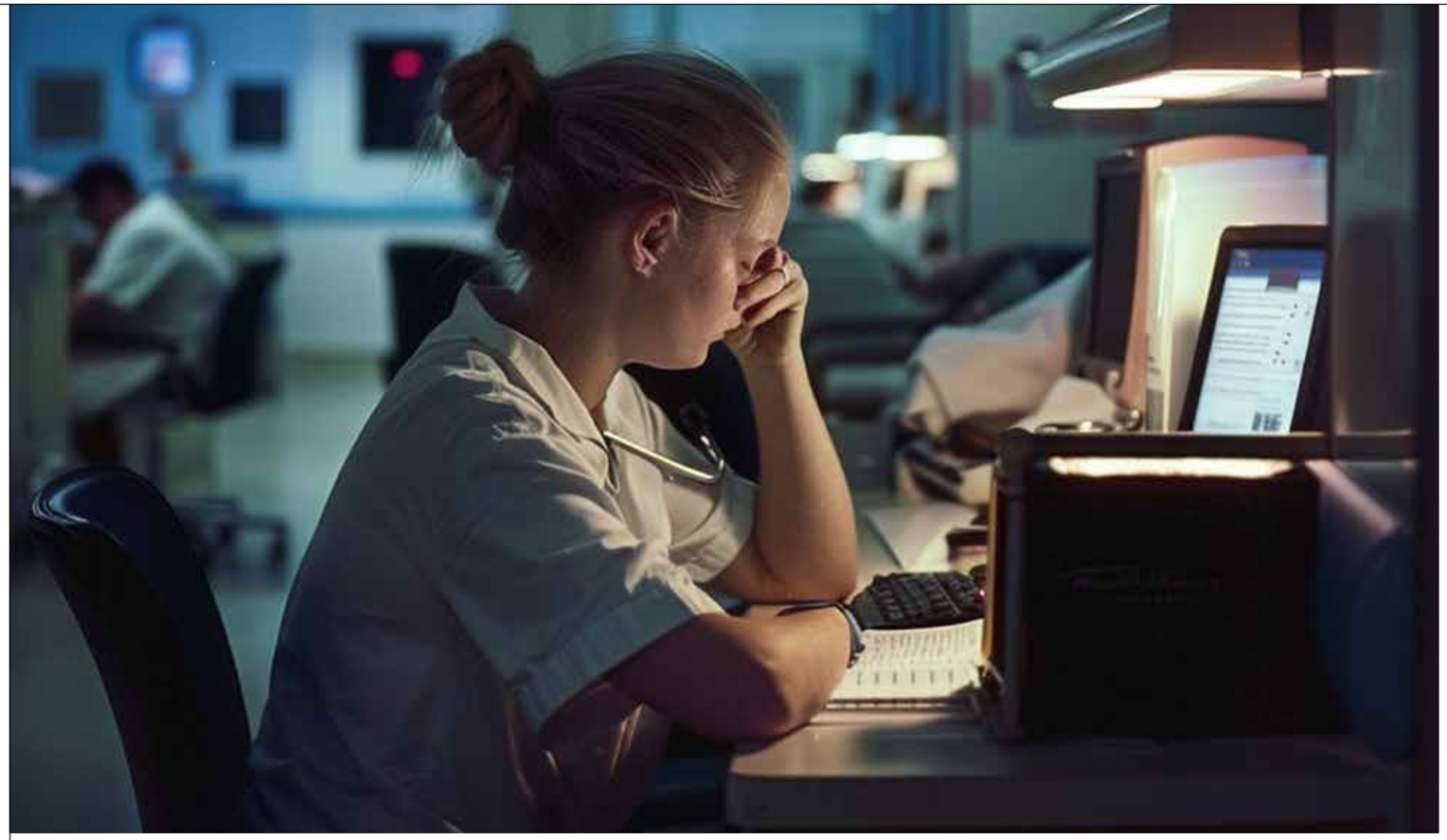

RAISE THE BARRIERS

Chris Handscomb *rethinks AI and the myth of proactive defence*

Threat intelligence is not a new concept. What has changed over the last decade is not the nature of cyber threats themselves, but the scale at which they operate and the widening gap between what organisations can theoretically see and what they can practically act upon. Much of the cybersecurity industry has blurred this distinction, marketing faster detection and richer forensic analysis as 'proactive' security. It is not. True proactive security is preventative by design; it exists to block hostile activity before it ever reaches or terminates inside the network, rather than helping teams investigate what has already gone wrong.

The uncomfortable reality is that cybersecurity is not facing a world of novel, exotic threats. It is facing the same threats, built on the same infrastructure, turned up to industrial scale. Networks are still networks and protocols are still protocols. Vulnerabilities are still overwhelmingly known rather than unknown. What has changed is how quickly those vulnerabilities can be discovered, weaponised and exploited, and how cheaply attackers can operate at volume.

From a security perspective, critical systems run on legacy platforms long past their official support life should not exist

Artificial intelligence sits at the centre of today's debate because it amplifies all of this. AI does not fundamentally change the mechanics of cyber attack or defence, rather it accelerates them. The previously human-led effort of researching a weakness, developing an exploit, building a botnet, conducting reconnaissance and deploying traffic against a target has not evolved into something new. It has simply been compressed. A process that once took weeks or months can now be executed in days or even hours. The method remains the same; the throttle has been opened.

This acceleration is often misinterpreted as increased sophistication, but in reality, it's increased volume. On any given day, a vulnerability in a popular plugin or application can be identified early in the morning, packaged into exploit kits before lunchtime and exploited at scale before the end of the working day. This is why most successful attacks still rely on unpatched, poorly monitored or misconfigured systems rather than so-called zero-day vulnerabilities. AI lowers the barrier to entry and multiplies attempts, but it does not transform the fundamentals.

The consequence of this shift is noise. Organisations find themselves buried under alerts, logs and signals that are technically correct, but operationally paralysing. Security teams normalise the background din, just as people acclimatise to constant noise in a busy city. Inevitably, the signal that matters is missed. In cybersecurity, the challenge is not seeing everything; it is knowing what to ignore so that the one indicator that matters stands out.

This is where the industry's growing reliance on reactive, AI-driven security creates a dangerous illusion of progress. Tools that promise intelligent detection, automated triage and AI-assisted response undoubtedly improve analyst productivity. But they intervene after the attacker has already arrived. If an AI system can observe malicious behaviour operating inside your environment, the adversary has succeeded in crossing the threshold and at that point prevention has failed. What remains is damage limitation.

There is an increasing tendency to suggest that AI can not only detect intrusions but also decide how to remediate them. This is where theory drifts sharply away from operational reality. Remediation is rarely a purely technical process, rather it is a judgement call informed by sector-specific risk, operational context, organisational history and human consequence. An AI agent may determine that the 'best' action is to patch, reconfigure or reboot a system. What it cannot fully understand is what that action means in the real world.

Consider healthcare environments, where critical systems regularly run on legacy, Windows-derived platforms long past their official support life. From a security textbook perspective, these systems should not exist. From a human perspective, they are keeping people alive. Does the AI prioritise eliminating the vulnerability or sustaining life? Who defines that hierarchy and who carries responsibility when something goes wrong?

The same dilemma extends to energy, transport and industrial environments, where changes in digital systems can have physical consequences. Applying a patch in the middle of live operations may technically improve security posture while simultaneously triggering outages, safety incidents or domino failures. Even in more conventional enterprise networks, professionals can point to countless examples where minor code changes or configuration updates have wiped databases, corrupted records or broken dependencies that had accumulated over decades. These are not edge cases; they are routine realities that experienced practitioners navigate precisely because they have lived through the consequences before.

AI, by contrast, infers its decisions from published knowledge, certification frameworks and academic literature. Systems such as Claude Mythos iterate over existing CVE data and research more rapidly than any human team could. That speed is valuable, but it is not wisdom. Nothing fundamentally new is being discovered; it is simply being aggregated and expressed more efficiently. Mistaking inference for understanding is the category error at the heart of much AI hype.

THE MORE DIVERSE AND HIGH QUALITY THE DATA, THE STRONGER THE PREDICTIVE VALUE

This matters because humans are predisposed to trust systems that communicate fluently. The tendency to project human traits, authority and even empathy onto machines has been documented since the ELIZA experiments in the Sixties, and it is resurfacing at scale today.

Cases of chatbot-induced psychosis and documented deaths linked to excessive dependence on conversational AI may sit outside traditional cybersecurity discussions, but they illustrate the same flaw. When we confuse output with comprehension, we outsource judgement to systems that cannot bear that responsibility.

Used correctly, AI is an extraordinary enabler. It can enrich data, add context at scale and dramatically improve analyst productivity. It can surface patterns that humans would miss under pressure and accelerate research that would otherwise take teams weeks. In this role, AI makes good humans better. What it cannot do is replace experienced humans.

Nowhere is this more apparent than within environments formed by legacy constraints. AI systems trained on generic security certifications and best-practice frameworks will almost always converge on the same conclusion: replace obsolete infrastructure. Practitioners know that this answer, while technically neat, often collides with reality. Systems cannot always be replaced quickly, cheaply or safely. Risk must be managed, not idealised away. Making those decisions requires contextual judgement, institutional memory and accountability – all of which remain irreducibly human.

This is the foundation of our approach to threat intelligence. Intelligence in itself is not the answer; it is a clue. Its value lies entirely in whether it can be acted upon in a way that meaningfully reduces risk. Monitoring indicators without context is little more than numerology. Understanding where hostile

infrastructure has been seen before, how it behaves across different sectors and what that behaviour means for a specific organisation is what turns information into protection.

True proactive threat intelligence accepts that vulnerabilities will exist. Rather than attempting to remediate everything after exposure, it focuses on denying attackers the opportunity to exercise those weaknesses in the first place. By blocking known hostile infrastructure, suppressing reconnaissance traffic and reducing background noise, organisations buy themselves time. This time is not wasted; it is used to prioritise remediation where it actually matters, rather than reacting under duress.

MOST ATTACKS STILL RELY ON UNPATCHED, POORLY MONITORED OR MISCONFIGURED SYSTEMS

Relying on a single vendor feed is like watching two pixels on a 4K screen and trying to call the score. Individual feeds overlap by a tiny percentage; they also conflict and often aren't tuned to your sector. Aggregation, plus adjudication, plus action is what turns clues into true protection.

The more diverse and high quality the data, the stronger the predictive value. When analysing this data, patterns begin to emerge and correlations can be drawn between seemingly unrelated events. For example, requests from specific geographies, unusual authentication attempts, or data flows that don't fit historical patterns can be prioritised intelligently. By doing this, we can see things before they can become destructive, for example, the MoveIT and Log4j vulnerabilities were seen as far out as 90 days before they became a problem for some.

The implications are particularly stark for critical national infrastructure. The most serious risks facing CNI today are not unknown zero-days, but well-documented weaknesses combined with extreme traffic volume and AI-enabled amplification. Public depictions often downplay this reality, framing AI as an abstract future concern rather than an accelerant of existing problems. Managing transparency and public confidence as well as strategic advantage is not easy, but pretending the problem is novel misses the point. It is not new; it is louder.

Within the cyber security industry, a divergence is now emerging between two philosophies. One emphasises AI-driven vulnerability discovery and ever faster research. The other prioritises prevention-first security that assumes flaws will persist and focuses on blocking hostile traffic before it causes harm. But speed without control is not security, and faster discovery alone does not prevent outages, breaches or disruption.

True proactive threat intelligence is quiet by design. When it works, nothing happens. Attacks fail upstream and background noise recedes. The one signal that matters becomes visible again. Monday-morning firefights turn into Monday-lunchtime summaries. Once organisations experience this shift, there is little appetite to return to chasing alerts after the fact.

AI will continue to evolve, and it will undoubtedly become more capable. It may replicate or accelerate parts of the intelligence workflow, from discovery to correlation. But replacing contextual judgement, accountability and human consequence is a far higher bar. Anyone claiming certainty about when, or whether, that bar will be cleared is being optimistic at best and irresponsible at worst.

The future of effective cybersecurity is not reactive, and it is not purely automated. It is preventative, human-guided and accountable. AI has a vital role to play in that future, but it belongs as an amplifier of human expertise, not a surrogate for it. ●

Chris Handscomb is Director of Solutions at Centripetal.

The growing reliance on reactive, AI-driven security creates a dangerous illusion of progress

Picture credit: AdobeStock