WHITEPAPER · 2026

Vision that *strikes* first.

AI-Augmented Offensive Security: A Framework for Continuous Adversarial Assessment in the Modern Enterprise.

SAQR AI · ABU DHABI
A BILINGUAL BRIEF FOR SECURITY LEADERS

رؤيةٌ تَستبقُ الضربةَ.

الأمن الهجومي المُعزَّز بالذكاء الاصطناعي: إطارٌ للتقييم العدائي المستمر في المؤسسات الحديثة.

صقر للذكاء الاصطناعي · أبوظبي
موجزٌ ثنائي اللغة لقادة الأمن السيبراني

القسم الأول

PART ONE · ARABIC

— الملخص التنفيذي

الأمن الهجومي في عصر الذكاء الاصطناعي

تغيّرت طبيعة التهديد السيبراني تغيُّراً جوهرياً خلال العامين الماضيين. لم يعد الخصم إنساناً يعمل وحده أمام شاشته، بل أصبح وكيلاً ذكياً قادراً على المسح والاستطلاع وصياغة الاستغلال وتنفيذه على نطاقٍ يفوق قدرات الفِرق الدفاعية التقليدية. في هذا الواقع الجديد، لم تَعُد دورات اختبار الاختراق السنوية أو نصف السنوية كافيةً لحماية المؤسسات الكبرى.

تأسست شركة "صقر للذكاء الاصطناعي" في أبوظبي لتقديم ردٍّ مكافئ: عملياتٌ هجومية مستمرة مدعومة بوكلاء ذكاء اصطناعي متقدمين، يعملون جنباً إلى جنب مع فريقٍ أحمر بشري من ذوي الخبرة الإقليمية العميقة. هذه الورقة تشرح إطار العمل الذي نقدّمه لقادة الأمن السيبراني، وتُبرز كيف يمكن للمؤسسات في دولة الإمارات والخليج العربي تبنّي هذا النهج لمواجهة جيلٍ جديد من التهديدات.

> "الصقر لا يَحرسُ، بل يَصطاد. يرى ما لا يراه غيرُه من ارتفاعٍ لا يبلغُه سواه، ويَنقَضُّ قبل أن تَعلمَ الفريسةُ أنها رُصِدت."

٢٤٠	×١٤	٩٨٪	٢٤/٧
ثغرة في المتوسط لكل مهمة	أسرع من الاختبار التقليدي	ثغرات حرجة قبل التصحيح	مراقبة عدائية مستمرة

المشهد المتغيّر للتهديدات

شَهِدت السنوات الأخيرة تحوّلاً ملحوظاً في طبيعة الجهات المهاجِمة وقدراتها. تستخدم مجموعات التهديد المتقدم المستمر نماذج لغوية كبيرة لتسريع مرحلة الاستطلاع، وصياغة رسائل تصيّد احتيالي مُقنِعة بلهجاتٍ محلية، وحتى توليد شيفرات استغلال مخصصة لبيئاتٍ بعينها. ما كان يستغرق فريقاً من المهاجمين أسابيع لإنجازه، أصبح ينفّذه وكيلٌ آلي في ساعات.

على الجانب الدفاعي، تواجه المؤسسات تحدياتٍ متراكمة: اتساع سطح الهجوم بفعل الانتقال إلى السحابة، انتشار البنية التحتية الظلّية، تعقُّد سلاسل التوريد الرقمية، وظهور أنظمة الذكاء الاصطناعي التوليدي كسطح هجومٍ جديد لم يكن موجوداً قبل سنواتٍ قليلة. كل هذه العوامل تفرض إعادة النظر في الافتراضات الأساسية التي بُنيت عليها برامج اختبار الاختراق التقليدية.

أين تخفق النماذج التقليدية

تعتمد معظم المؤسسات اليوم على دورة تقييمٍ سنوية أو ربع سنوية، تنتهي بتقريرٍ مفصّل يصف ثغراتٍ تم اكتشافها في لحظةٍ زمنية معيّنة. في الوقت الذي يصل فيه التقرير إلى مكتب رئيس أمن المعلومات، يكون نصف النتائج قد تجاوزه الزمن، إمّا لأن الأنظمة تَغيّرت، أو لأن تهديداتٍ جديدة ظهرت لم تكن مشمولةً في نطاق التقييم الأصلي.

- **الفجوة الزمنية:** المتوسط بين دورتَي اختبار يتجاوز التسعين يوماً، وهي مدةٌ كافية لاستغلال أي ثغرةٍ حرجة عدّة مرات.
- **محدودية النطاق:** تُجرى الاختبارات عادةً على عيّناتٍ من الأصول، وليس على البنية التحتية الكاملة، ما يَترك مناطقَ مظلمة دون تغطية.
- **غياب البُعد التكيُّفي:** المختبِر البشري، مهما بلغت كفاءته، لا يستطيع محاكاة خصمٍ يعمل على مدار الساعة بقدرات حوسبةٍ عالية.
- **تجاهل أنظمة الذكاء الاصطناعي:** القليل من برامج التقييم تتناول مخاطر النماذج اللغوية وأنظمة الوكلاء الذكية المنشورة داخلياً.

إطار العمل: التقييم العدائي المستمر

يَقوم إطار العمل الذي نقدّمه على أربعة أعمدة أساسية تعمل بشكلٍ متكامل لإنشاء قدرةٍ هجومية دائمة، لا مجرد فعالية موسمية.

أولاً: الاستطلاع المُؤتمَت

يَعمل وكلاء صقر باستمرارٍ لرسم سطح الهجوم الكامل للمؤسسة: النطاقات الفرعية، الخدمات المكشوفة، تكوينات السحابة الخاطئة، التسريبات في مستودعات الشيفرة العامة، البيانات المسرّبة في الويب المظلم. هذا الرسم يتم تحديثه بشكلٍ مستمر، لا في لحظاتٍ منعزلة، وهو ما يُتيح لقيادة الأمن رؤيةً صادقة لما يراه الخصم لحظةً بلحظة.

ثانياً: المحاكاة العدائية المتقدمة

تتولّى وكلاء ذكاء اصطناعي متخصصة تنفيذ سلاسل استغلالٍ مُعقّدة ضد البيئات المستهدفة، بالاعتماد على معرفةٍ محدّثة بالثغرات والتقنيات. الفريق البشري يُشرف على هذه العمليات ويُوجّهها نحو السيناريوهات الأكثر صلةً بالمؤسسة، خاصةً تلك المرتبطة بالمشهد التهديدي الإقليمي.

ثالثاً: تقييم أنظمة الذكاء الاصطناعي

مع تبنّي المؤسسات الإماراتية والخليجية أنظمة الذكاء الاصطناعي التوليدي بسرعةٍ متزايدة، أصبح تأمين هذه الأنظمة ضرورةً ملحّة. نُجري اختباراتٍ لمقاومة كسر الحماية، وتسريب البيانات عبر الحقن، وتجاوز حدود الوكلاء الذكية، وأمن سلسلة التوريد للنماذج.

رابعاً: التنسيق مع الفريق الأزرق

لا قيمةَ لاكتشاف الثغرات دون قدرةٍ على تحويلها إلى تحسيناتٍ دفاعية ملموسة. يَعمل فريقنا بصورةٍ تَكاملية مع فرق العمليات الدفاعية لدى العملاء، في نموذج "الفريق الأرجواني"، لضمان أن كل اكتشافٍ يُترجَم إلى إجراءٍ وقائي قابلٍ للقياس.

الإطار التنظيمي الإماراتي

تَتبنّى دولة الإمارات أحد أكثر الأُطر التنظيمية تقدُّماً في المنطقة في مجال الأمن السيبراني. يَفرض المجلس الأعلى للأمن السيبراني، والهيئات القطاعية كهيئة الاتصالات والحكومة الرقمية، متطلباتٍ صارمة على المؤسسات الحيوية والمصرفية والطاقة. إطار عملنا مصمَّمٌ ليتوافق مع هذه المتطلبات، ويتجاوزها، عبر تقديم أدلةٍ مستمرة على الامتثال بدلاً من لقطاتٍ سنوية.

- توافقٌ كامل مع معايير الهيئة الوطنية للأمن السيبراني (سابقاً NESA).
- دعمٌ لمتطلبات مصرف الإمارات المركزي للمؤسسات المالية.
- قدرةٌ على تقديم تقارير بصيغٍ متوافقة مع لوائح حماية البيانات الشخصية.
- تخزينٌ آمن للبيانات الحساسة داخل الحدود السيادية للدولة.

منهجية العمل

تَبدأ كل مهمةٍ بجلسة تأطيرٍ مع قيادة الأمن لدى العميل، تُحدَّد فيها الأصول الحرجة، نموذج التهديد، والقيود التشغيلية. ثم يُنشر بنيةٌ تحتية مخصصة للعميل تَعمل بمعزلٍ تامّ عن مهام العملاء الآخرين. تَستغرق الدورة الأولى عادةً ما بين ٧٢ ساعةً وأسبوعَين، حسب حجم سطح الهجوم.

بعد الدورة التأسيسية، تنتقل المؤسسة إلى نمطٍ مستمر يَستمر فيه الاستطلاع والمحاكاة على مدار العام، مع جلسات استعراضٍ شهرية يُقدَّم فيها أبرز الاكتشافات، وأولوياتها، والمسارات الموصى بها للمعالجة. هذا النموذج يَستبدل التقرير السنوي السميك بحوارٍ مستمر بين فريقنا وقيادة الأمن.

دراسة حالة موجزة

تَعاوَنّا مؤخراً مع إحدى المؤسسات المالية الكبرى في أبوظبي، التي كانت تُجري اختبار اختراقٍ سنوياً من خلال مزوّدٍ دولي. خلال الأسبوعين الأولين من العمل معنا، اكتشف وكلاؤنا ثلاثَ ثغراتٍ حرجة لم تظهر في آخر تقريرٍ سنوي للعميل، إحداها كانت قد ظهرت قبل أربعة أشهر فقط نتيجة تحديثٍ في بيئة الاختبار. هذه الفجوة الزمنية تَختصِر القصة كاملةً: ما لا يُختبَر باستمرار، يُستغَلّ.

الخلاصة والدعوة للعمل

الأمن الهجومي ليس ترفاً تنظيمياً، بل ضرورةٌ استراتيجية لكل مؤسسةٍ جادّة في حماية أصولها وعملائها. الفارق بين المؤسسة التي تُدرك ثغراتها قبل خصمها، والمؤسسة التي تَكتشفها بعد فوات الأوان، هو فارقٌ يُقاس بمئات الملايين من الدراهم وبسمعةٍ يصعب استعادتها.

نَدعو قادة الأمن في المؤسسات الإماراتية والخليجية إلى التواصل معنا لإجراء تقييم سطح هجومٍ مبدئي دون التزام. النتائج تُسلَّم خلال ٧٢ ساعة، وتَكفي وحدها لتُغيّر الصورة التي تَملِكها قيادتُك عن وضعها الأمني الحقيقي.

Part Two

PART TWO · ENGLISH

— EXECUTIVE SUMMARY

Offensive Security in the *AI Era*

The nature of cyber threat has shifted fundamentally over the past two years. The adversary is no longer a lone operator at a keyboard. It is an autonomous agent capable of reconnaissance, exploit synthesis, and execution at a scale that traditional defensive teams cannot match. In this new reality, annual or semi-annual pentest cycles are no longer sufficient to defend modern enterprises.

Saqr AI was founded in Abu Dhabi to provide a symmetric response: continuous offensive operations driven by advanced AI agents, working alongside a human red team with deep regional expertise. This paper outlines the framework we offer security leaders, and how organisations in the UAE and wider Gulf can adopt this approach to face a new generation of threats.

> *"The falcon does not guard. It hunts. It sees what others cannot from a height others cannot reach, and it strikes before the prey knows it has been seen."*

240	14×	98%	24/7
CVES / ENGAGEMENT	FASTER THAN PENTESTS	CRITICAL PRE-PATCH	ADVERSARIAL COVERAGE

The Shifting Threat Landscape

Recent years have seen a marked shift in attacker capability. Advanced persistent threat groups now leverage large language models to accelerate reconnaissance, craft locally-fluent phishing in regional dialects, and even generate exploit code tailored to specific environments. What once took a team of attackers weeks now executes in hours via autonomous agents.

On the defensive side, enterprises face compounding challenges: expanding attack surface from cloud migration, proliferating shadow infrastructure, growing complexity of digital supply chains, and the emergence of generative AI systems as a new attack surface that did not exist a few years ago. Each of these forces calls into question the assumptions underlying traditional pentest programmes.

Where the Traditional Model Breaks

Most organisations today rely on annual or quarterly assessment cycles that conclude with a detailed report describing vulnerabilities discovered at a specific point in time. By the time the

report reaches the CISO's desk, half the findings are already stale — either the systems have changed, or new threats have emerged that fell outside the original scope.

- **The temporal gap:** the average interval between assessment cycles exceeds 90 days, a window long enough for any critical vulnerability to be exploited several times over.
- **Limited scope:** assessments are typically run on asset samples, not full infrastructure, leaving dark zones uncovered.
- **No adaptive dimension:** a human tester, however skilled, cannot simulate an adversary operating around the clock with significant compute behind them.
- **AI systems ignored:** few assessment programmes address the risks of LLMs and agentic systems deployed inside the enterprise.

The Framework: Continuous Adversarial Assessment

The framework Saqr AI proposes rests on four pillars working in concert to create a permanent offensive capability, not a seasonal event.

1. Autonomous Reconnaissance

Saqr's agents continuously map the enterprise's full attack surface: subdomains, exposed services, cloud misconfigurations, leaks in public code repositories, credentials surfacing on the dark web. The map updates continuously, not in isolated moments, giving security leadership an honest view of what an adversary sees minute by minute.

2. Advanced Adversarial Simulation

Specialised AI agents execute complex exploit chains against target environments, drawing on up-to-date knowledge of vulnerabilities and techniques. Human operators oversee these operations and steer them toward scenarios most relevant to the client, particularly those tied to the regional threat landscape.

3. AI System Assessment

As UAE and Gulf enterprises adopt generative AI systems at pace, securing those systems has become an urgent need. We run jailbreak resistance testing, indirect prompt injection assessments, agent boundary testing, and model supply chain security review.

4. Purple Team Integration

Finding vulnerabilities has no value without the ability to translate them into concrete defensive improvements. Our team works in close integration with client blue teams in a purple team model, ensuring every finding maps to a measurable preventative action.

The UAE Regulatory Frame

The UAE maintains one of the most advanced cybersecurity regulatory frameworks in the region. The UAE Cybersecurity Council, along with sector authorities such as the TDRA, imposes strict requirements on critical, financial, and energy sector entities. Our framework is designed to meet these requirements, and exceed them, by providing continuous evidence of compliance rather than annual snapshots.

- Full alignment with UAE Information Assurance standards (formerly NESA).
- Support for Central Bank of the UAE requirements for financial institutions.
- Reporting formats compatible with UAE Personal Data Protection regulations.
- Secure storage of sensitive client data within the UAE's sovereign boundaries.

Engagement Methodology

Every engagement begins with a framing session between Saqr AI and the client's security leadership. Critical assets, threat model, and operational constraints are defined together. A client-dedicated infrastructure is then provisioned, isolated entirely from other client engagements. The initial cycle typically takes between 72 hours and two weeks, depending on attack surface size.

After the foundation cycle, the organisation transitions into a continuous mode in which reconnaissance and simulation run throughout the year, with monthly review sessions presenting the most significant findings, their priority, and recommended remediation paths. This model replaces the thick annual report with an ongoing dialogue between our team and the client's security leadership.

Brief Case Study

We recently worked with a major Abu Dhabi-based financial institution that had been running annual pentests through an international provider. Within the first two weeks of engagement with Saqr AI, our agents surfaced three critical vulnerabilities not present in the client's most recent annual report — one of which had appeared only four months earlier as a result of a test environment update. The temporal gap tells the entire story: what is not continuously tested is exploited.

Conclusion and Next Steps

Offensive security is not an organisational luxury. It is a strategic necessity for any enterprise serious about protecting its assets and customers. The difference between the organisation that

knows its vulnerabilities before its adversary does, and the one that finds out after the fact, is measured in hundreds of millions of dirhams and reputational damage that takes years to repair.

Saqr AI invites security leaders across UAE and Gulf enterprises to engage us for an initial, no-obligation attack surface assessment. Results are delivered within 72 hours, and on their own they tend to change the picture leadership holds of its true security posture.

صقرٌ لا يَحرس.
يَصطاد.

A falcon does not guard.
It hunts.

SAQR AI
ABU DHABI GLOBAL MARKET
HELLO@SAQR.AI · +971 2 XXX XXXX

صقر للذكاء الاصطناعي
سوق أبوظبي العالمي
hello@saqr.ai · +٩٧١٢