



2026/10/26

الذكاء الاصطناعي أداة للتسليح

قراءة تحليلية في ضوء تقرير شركة Anthropic بشأن إساعة
استخدام الذكاء الاصطناعي

عمر فراس الزبيدي

● مقال رأي

الذكاء الاصطناعي أداة للتسليح: قراءة تحليلية في ضوء تقرير شركة Anthropic بشأن إساءة استخدام الذكاء الاصطناعي

سلسلة إصدارات مركز البيان للدراسات والتخطيط / قسم الأبحاث / الدراسات السياسية

الصادر / مقال رأي

الموضوع / شؤون إقليمية ودولية / مكافحة التطرف والإرهاب

عمر فراس الزبيدي / باحث

عن المركز

مركزُ البيان للدراسات والتخطيط مركزٌ مستقلٌّ، غيرٌ ربحيٍّ، مقرُّه الرئيس في بغداد، مهمته الرئيسية -فضلاً عن قضايا أخرى- تقديم وجهة نظر ذات مصداقية حول قضايا السياسات العامة والخارجية التي تخصُّ العراق بنحو خاص، ومنطقة الشرق الأوسط بنحو عام. ويسعى المركز إلى إجراء تحليلٍ مستقلٍّ، وإيجاد حلول عملية جُلِّتْ لقضايا معقدة تهتمُّ الحقلين السياسي والأكاديمي.

ملحوظة:

لا تعبّر الآراء الواردة في المقال بالضرورة عن اتجاهات يتبناها المركز، وإنّما تعبّر عن رأي كاتبها.

حقوق النشر محفوظة © 2026

www.bayancenter.org

info@bayancenter.org

Since 2014

المقدمة

شهدت التقييمات الأمنية المتعلقة بتوظيف النماذج اللغوية الكبيرة في الأغراض العدائية، حتى وقت قريب، تركيزاً على السيناريوهات الافتراضية والتجارب المختبرية المقيدة، من دون الاستناد بصورة كافية إلى معطيات مستمدة من عمليات فعلية. وفي هذا السياق، يقدم تقرير الاستخبارات التهديدية الصادر عن شركة أنثروبك (Anthropic) في أيلول/سبتمبر 2026 معطيات تجريبية تسهم في إعادة توجيه مسار التقييم الأمني، إذ توثق العمليات المرصودة خلال الفترة الممتدة من كانون الأول/ديسمبر 2025 إلى آب/أغسطس 2026، وفي بيئات جغرافية وتنظيمية متباينة، انتقالاً نحو توظيف الذكاء الاصطناعي التوليدي في تطبيقات عدائية عملية، بما يعكس تنامي دوره بوصفه أداة يمكن تسخيرها في أنشطة تهدد الأمن السيبراني والأمن الدولي.

وتكشف السجلات عن محاولات لتوظيف هذه التقنيات في تطوير أسلحة تقليدية وبيولوجية، وتنفيذ عمليات سيبرانية، ودعم أنشطة ذات أغراض عدائية متنوعة. وقد توزعت الأنشطة المرصودة على سبع مسارات رئيسية، هي: الهجمات الإلكترونية، وعمليات التأثير، والمراقبة، والاحتلال، والانتهاكات البيولوجية، وتطوير الأسلحة التقليدية، والتقطير غير المشروع لقدرات النماذج (Model Distillation). ويُقصد بالتقطير، في هذا السياق، استغلال مخرجات النماذج اللغوية المتقدمة في تدريب نماذج أصغر حجماً وأقل كلفة، بما يتيح نقل بعض قدراتها إلى أنظمة أخرى، ويثير تساؤلات بشأن حدود الاستخدام المشروع لهذه التقنية وإمكانات توظيفها في تجاوز القيود المفروضة على الوصول إلى النماذج المتقدمة.

وتشير المعطيات الواردة في التقرير إلى أن النماذج التي رُصد استغلالها تنتمي إلى إصدارات Claude، وهي Haiku و Sonnet و Opus. وفي المقابل، لم يوثق التقرير مؤشرات على تورط نماذج الجيل المتقدم، مثل Fable أو Mythos، في هذه الأنشطة، باستثناء حادثة تقطير واحدة. وتبرز أهمية هذه المعطيات في أنها تنقل النقاش بشأن الاستخدام العسكري والهجومي لأدوات الذكاء الاصطناعي من نطاق الاحتمالات النظرية إلى دراسة أنماط الاستخدام الفعلي، وما يرتبط بها من تحديات تقنية وأمنية وتنظيمية.

وانطلاقاً من ذلك، لم يعد التحدي مقتصرًا على إثبات قابلية النماذج اللغوية الكبيرة للاستخدام في الأغراض العدائية، بل بات يشمل تطوير آليات منهجية لرصد هذا الاستخدام والحد من انتشاره واحتواء آثاره الأمنية المباشرة وغير المباشرة. ويقتضي هذا التحول إعادة النظر في أولويات البحث الأمني، وتعزيز التنسيق بين الجهات المعنية، وتطوير أدوات حوكمة تستجيب للطبيعة المتغيرة للتهديدات المرتبطة بالذكاء الاصطناعي التوليدي.

وفي هذا الإطار، تبحث الورقة مدى قدرة أطر حظر التسلح التقليدية على استيعاب التهديدات غير المادية الناشئة عن امتلاك المعرفة البرمجية وتداولها وتوظيفها في الأغراض العدائية. كما تسعى إلى استكشاف إمكانات تطوير معايير أمنية جماعية تتجاوز الاعتماد الحصري على بروتوكولات الرقابة السحابية وآليات التصفية التجارية، باتجاه بناء إطار استراتيجي أكثر شمولاً يتلاءم مع الطبيعة اللامركزية للتسلح الخوارزمي، ويوازن بين متطلبات الابتكار التقني وضرورات الأمن الدولي والحد من مخاطر الانتشار.

حجم المشكلة وتوظيف الذكاء الاصطناعي في التسليح

تشير المعطيات التوثيقية في تقرير أنثروبك إلى نشاط خلية عملياتية في شمال اليمن تدير ثلاثة برامج صاروخية متزامنة.

حيث يركّز المسار الأول على مقذوف موجّه يعتمد على حاسوب طيران مشتقّ من تقنيات الهواتف الذكية بوصفه جهازاً مركزياً، ومقترناً بنظام توجيه نهائي. ويستهدف المسار الثاني تطوير منظومة باليستية متعددة المراحل بمدى يتجاوز 2000 كيلومتر، فيما يركّز المسار الثالث على سلسلة صواريخ (R2000)، التي تتضمن تصميماً لمركبة انزلاقية فرط صوتية (HGV)، وهي نظام تسليح متقدم ينطلق بسرعات تتجاوز خمسة أضعاف سرعة الصوت) ماخ 5 فما فوق). وتُحمل هذه المركبات في مرحلة الإطلاق الأولى بواسطة صواريخ باليستية تقليدية، لتبلغ طبقات الغلاف الجوي العلوي، ثم تنفصل عن الصاروخ الناقل وتبدأ مساراً انزلاقياً مستقلاً نحو الهدف دون الخروج إلى الفضاء الخارجي، حيث تنحصر هذه التقنيات عادةً ضمن ترسانات القوى العسكرية الكبرى.

وتكمن الأهمية التحليلية لهذه الحالة في الاستعاضة الهيكلية عن الطواقم الهندسية البشرية بمنظومة «Claude Code»، حيث تُكَلَّف النماذج اللغوية بمهام تصميم برمجيات التحكم، وخوارزميات تقدير الموقع، ودمج أنظمة الطيران الآلي مفتوحة المصدر مع العتاد الصلب. كما أُسندت إلى النماذج مهام ضبط معاملات التحكم، وإدارة بيئة بناء البرمجيات الثابتة (Firmware)، وتنفيذ عمليات المحاكاة الأولية، فتحوّلت النماذج اللغوية هنا إلى بنية هندسية بديلة للمبرمجين بالكامل.

تُظهر السجلات التشغيلية استيعاباً دقيقاً لمعمارية وكلاء الذكاء الاصطناعي (AI Agents). وقد نُظِّمت بيئة العمل عبر تشغيل مثيلات متوازية من النموذج، خُصِّصت إحداها للبرمجة، والثانية للبحث التقييبي عن الأكواد، فيما تولّت الثالثة مهام التدقيق والمراجعة. ويُقارب هذا التقسيم الوظيفي ديناميكيات الفرق الهندسية المصغّرة.

وقد انعكست هذه الهيكلية بوضوح على تسريع دورة التطوير. وعقب إخفاق اختبار إطلاق أحد المقذوفات، جرى اللجوء إلى النموذج لتشخيص الأعطال وتحليل أسباب الفشل خلال

إطار زمني لا يتجاوز بضع ساعات. وبموازاة ذلك، طُوّرت الخلية أدوات محاكاة حركية مستقلة تعمل دون اتصال بشبكة الإنترنت، بمعزل عن بيئات الحوسبة التجارية، مثل «MATLAB»، أو خوادم الذكاء الاصطناعي.

ويؤسس هذا الاستقلال التشغيلي لقدرات هندسية مستدامة؛ إذ يُحتمل أن تستمر هذه البنى التطويرية الجوهرية في العمل بكفاءة، متجاوزة قيود الانقطاع، حتى في حال تفعيل بروتوكولات الحظر التقني السحابي ضد المستخدمين.

الحالات الدولية وحدود الاحتواء (روسيا وأوكرانيا والصين)

تشير المعطيات إلى أن النماذج التشغيلية المعتمدة على الذكاء الاصطناعي في التطوير العسكري تتجاوز النطاق الإقليمي للحالة اليمينية؛ إذ رُصدت في السجلات ست حالات إضافية لبرامج تطوير أسلحة تقليدية. وتتركز الأنماط الأكثر تعقيداً لهذه الأنشطة ضمن النطاقين الجغرافيين لروسيا والصين، بالتزامن مع توجيه عمليات استهداف سببراني نشطة نحو البنى التحتية الأوكرانية. وتُظهر البيانات الروسية مسارين متوازيين للاستغلال التقني. فعلى الصعيد المادي، انخرط فاعلون مستقلون في محاولتين منفصلتين لتطوير خوارزميات للتحكم والقيادة، مخصصة لأسراب الطائرات المسيّرة الانتحارية ذاتية التوجيه. أما على الصعيد الرقمي، فتشير المعطيات إلى احتمال ارتباط مجموعة تجسس متقدمة، تحمل مؤشرات تقنية تتطابق مع تلك المرتبطة بعمليات مجموعة (APT29)، وهي كيان سببراني هجومي متقدم (Advanced Persistent Threat) يُعتقد أنه يحظى برعاية دولة، بتنفيذ اختراقات منظّمة. وقد اعتمدت هذه الاختراقات على مسارات عمل مؤتمتة بالذكاء الاصطناعي لرفع كفاءة هجمات التصيد الاحتيالي وتوسيع نطاق استخراج البيانات.

واتجهت النسبة الأكبر من هذه الهجمات الروسية نحو القطاعات الحكومية والدبلوماسية في أوكرانيا والمنظومة الأوروبية. كما سُجّلت محاولات استهداف نوعية ركّزت تحديداً على

اختراق قواعد بيانات الملكية الفكرية للشركات المصنّعة للطائرات المسيّرة العاملة ضمن المسرح الأوكراني.

ويبرز المسرح الأوكراني بوصفه بيئة هندسية متقدمة لتطبيق ما يُعرف بـ “شبكات القتل الأفقية»، وهي هيكلية تعتمد على دمج خوارزميات الرؤية الحاسوبية، عبر منصات مخصصة، مع شبكات الطائرات المسيّرة لتحليل التغذية البصرية لحظياً. وبمجرد رصد الآليات العسكرية، يتولى وكيل الذكاء الاصطناعي تصنيفها تلقائياً وإرسال إحداثياتها مباشرة إلى طبقات البيانات المشتركة، مما يقلّص المدة الزمنية الفاصلة بين الرصد والاستهداف، متجاوزاً التراتبية القيادية. ويعرّز هذا النمط التشغيلي قدرة المقذوفات على التتبع النهائي المستقل للأهداف، مما يحدّ جزئياً من فاعلية التشويش الإلكتروني القائم على قطع الاتصال.

تأخذ المؤشرات المرصودة في الصين طابعاً هندسياً وتنظيمياً متبايناً حيث تشير ثلاث حالات موثقة إلى إدماج “Claude Code” مباشرة في صياغة وهندسة أنظمة التوجيه والتحكم الخاصة بمنصات التسليح التقليدية حيث يظهر ارتباط النماذج اللغوية بتضخيم قدرات الأطراف الفاعلة غير الدولية عبر رصد حملة سيبرانية انطلقت من مقاطعة «هونان» ونُفذت بواسطة طلاب جامعيين و تشير التحليلات التشغيلية إلى توظيف هؤلاء الأفراد للنموذج كطبقة هندسية وتنسيقية عليا لإدارة حزم هجومية عابرة للحدود و شمل بنك الأهداف شبكات حكومية وتجارية موزعة جغرافياً عبر الشرق الأوسط، أوروبا، وجنوب شرق آسيا.

التحديات في المجال البيولوجي

تشير السجلات التشغيلية إلى خمس حوادث موثقة سُخِّرَ فيها النموذج «Claude» لخدمة استقصاءات علمية تظهر ارتباطاً محتملاً بتطوير عوامل بيولوجية خطيرة تتضمن أبرز هذه الوقائع نشاطاً بحثياً موجهاً لصالح معهد عسكري حيث تركزت الاستفسارات حول هندسة وتطبيق بروتوكولات «اكتساب الوظيفة» (Gain-of-Function) على فيروس شيكونغونيا (Chikungunya virus). تشير المعطيات التقنية للشركة إلى تدخل بروتوكولات الأمان الداخلي لتحجيم هذا المسار الاستقصائي أدت آليات الرقابة الآلية إلى تخفيض مستوى الدعم المعرفي المتاح وحصر تفاعلات المستخدم قسراً ضمن إصدارات النماذج ذات القدرات التحليلية الأدنى مما قلّص حجم المخرجات القابلة للتطبيق العملي المباشر.

تطور الذكاء الاصطناعي في الحروب والتصنيع العسكري ودعم الإرهاب

تتجاوز التداعيات الاستراتيجية المترتبة على إدماج تقنيات الذكاء الاصطناعي في المنظومات الأمنية والعسكرية حدود الحوادث الفردية أو الاستخدامات المعزولة. تشير البيانات التحليلية إلى تحولات هيكلية عميقة تعيد تشكيل طبيعة حروب المواجهة التقليدية وتمتد لتشمل سلاسل الإنتاج والتصنيع العسكري المتقدم. يتيح هذا المسار التكنولوجي للتنظيمات المسلحة والفاعلين من غير الدول مساحات حركة وتطوير غير مسبقة حيث يكمن التهديد الجوهري في قدرة هذه الخوارزميات على التخفيض الجذري للتكلفة المعرفية والتنظيمية التي كان يُشترط توافرها لإنشاء منظومات عسكرية معقدة ومن المحتمل أن يُسهم ذلك في تقليص الفجوة التقنية التاريخية التي طالما فصلت بين القدرات التسليحية للدول وتلك الخاصة بالفاعلين المستقلين.

أما فيما يتعلق بأنظمة الاستهداف وإدارة مسرح العمليات تعمل خوارزميات معالجة البيانات الضخمة على إعادة الهيكلة الجذرية لآليات اتخاذ القرار العسكري ف تُظهر النماذج التحليلية المتقدمة قدرة استثنائية على الدمج اللحظي والتزامني للصور الأقمار الفضائية ومعطيات الاستخبارات والبيانات الميدانية المستمدة من المستشعرات الأرضية و يُوظف هذا الدمج في توليد بنوك من الأهداف الدقيقة وإجراء تقييمات آلية فورية للأضرار يقلص «ضباب الحرب» تدريجياً , تُزود القيادات الميدانية بمعطيات تنتبأ لتقدير مسارات الخصم وتحركاته مما يوجه الحركة التكتيكية وإدارة خطوط الإمداد نحو قطاعات أكثر أماناً. وعلى صعيد الأسلحة ذاتية التشغيل والمسيرات الذكية يُطرح دمج شرائح المعالجة الطرفية (Edge AI) وتقنيات الرؤية الحاسوبية داخل الهياكل الجوية كحل هندسي لضمان استمرار العمليات داخل بيئات التشويش الإلكتروني الكثيفة و أن هذه البنية التقنية قد توفر استقلالية توجيهية وملاحية تامة للمقذوفات لتعمل بمعزل عن إشارات التموضع العالمي (GPS) أو الحاجة للارتباط بمحطات التحكم الأرضية و تتجه مسارات التطوير نحو بناء أسراب لامركزية؛ شبكات جوية قادرة على التواصل الذاتي لتبادل البيانات الميدانية وتوزيع المهام الحركية وتطويق منظومات الدفاع الجوي المعادية دون تدخل بشري مباشر.

يمتد الأثر التشغيلي لهذه البيانات ليصل إلى البنية التحتية للصناعات الدفاعية وسلاسل الإنتاج المادي عبر خوارزميات التصميم التوليدي (Generative Design). تقترح هذه المنظومات هياكل فيزيائية وتصاميم هندسية متقدمة للمقذوفات وأنظمة الدروع، بحيث تُحقق توازناً رياضياً دقيقاً بين تقليل الكتلة والحفاظ على الصلابة الهيكلية، إلى جانب تحسين الخصائص الأيروديناميكية بمستويات تتجاوز قدرات النمذجة اليدوية المعيارية. تدعم هذه المنظومة أيضاً بروتوكولات الصيانة التنبؤية. تبرز المؤشرات الإحصائية ارتباطاً وثيقاً بين المراقبة الخوارزمية المستمرة للآليات وبين ارتفاع معدلات الجاهزية القتالية، من

خلال تشخيص احتمالات الإخفاق الميكانيكي قبل وقوعها الفعلي، مما ينعكس إيجاباً على خفض الأعباء اللوجستية.

حيث تُظهر المعطيات الميدانية توجهاً متزايداً لدى التنظيمات المسلحة والفاعلين من غير الدول نحو استغلال النماذج الخوارزمية لخفض الحواجز المعرفية والتنظيمية المرتبطة بالتهديدات الأمنية. يبرز هذا التحول الهيكلي بوضوح في مسارات النمذجة الهندسية للمقذوفات البديلة. تشير السجلات إلى دمج نماذج الرؤية الحاسوبية مفتوحة المصدر ضمن هياكل الطائرات المسيّرة التجارية منخفضة التكلفة. يوفر هذا الدمج التقني استقلالية توجيهية تتيح للأنظمة التعرف التلقائي على الأهداف البشرية والآليات العسكرية، وتنفيذ مسارات الانقضاض النهائي بمعزل عن التوجيه البشري. تُبطل هذه المقاربة اللامركزية فاعلية أنظمة التشويش الإلكتروني المعيارية التي تعتمد على قطع روابط الاتصال اللاسلكي. تمتد التطبيقات الاستراتيجية لهذه الخوارزميات لتشمل إدارة العمليات النفسية وهندسة الاستقطاب. تشير التحليلات السلوكية إلى انتقال تدريجي نحو توظيف تقنيات التوليد الآلي والتزييف العميق (Deepfake) لإنتاج مواد مرئية وصوتية متقدمة الدقة، مصحوبة بمسارات ترجمة فورية، بهدف تمرير الرسائل الأيديولوجية الراديكالية. يترافق هذا النمط مع رصد نشر وكلاء استقلاليين (AI Agents) مبرمجين لإدارة شبكات واسعة من الحسابات الوهمية عبر منصات التواصل الاجتماعي. تتكيف هذه البرمجيات تفاعلياً مع السمات السيكلوجية للشرائح المستهدفة، مما يُعزز كفاءة عمليات التجنيد ونشر الفكر المتطرف. على صعيد التسلح غير التقليدي، تخرق النماذج اللغوية الكبيرة القيود المرتبطة بالتخصص الأكاديمي والمؤسسي. تشير البيانات الاستقصائية إلى فاعلية هذه المنظومات في تجاوز بروتوكولات الأمان الرقمية، وتوفير أدلة إجرائية دقيقة لتخليق المركبات الكيميائية السامة والعوامل البيولوجية. تقترح هذه النماذج الخوارزمية بدائل عملية

وهندسية لتعويض غياب المواد الأولية المحظورة تجارياً. يتيح هذا المسار لعناصر تقتصر إلى التكوين العلمي المنهجي تصميم وإنتاج تهديدات غير متماثلة ذات أثر استراتيجي.

يُمثل النطاق السيبراني مساراً موازياً لتأمين البنية التحتية واللوجستية لهذه التنظيمات. تُوظف قدرات المعالجة اللغوية المتقدمة في صياغة هجمات تصيد احتيالي (Spear-phishing) عالية التخصيص، تستهدف مسؤولي الشبكات الحساسة لانتزاع صلاحيات الوصول بشكل سري. من المحتمل أن تلعب الأكواد البرمجية المولدة آلياً دوراً مباشراً في تسريع وتيرة تطوير البرمجيات الخبيثة. تساهم هذه الأكواد في اكتشاف ثغرات أمنية غير موثقة (Zero-day vulnerabilities) وتطوير بروتوكولات تشفير معقدة تُستخدم في هجمات الفدية، مما يُؤسس لغطاء تمويلي مستدام يغذي العمليات الميدانية والأنشطة المحظورة.

تقييم دقة وموثوقية وكلاء الذكاء الاصطناعي في الأعمال العسكرية والتنسيق الهندسي

يطرح تقرير انثروبك إشكالية منهجية حول مستوى الموثوقية الفنية لوكلاء الذكاء الاصطناعي عند إدماجهم في مشاريع التسليح المعقدة من قبل الفاعلين غير الحكوميين، وتحديدًا في هندسة برامج التوجيه الصاروخي.

النماذج اللغوية التوليدية، بطبيعتها الاحتمالية، لا تُنتج مخرجات حتمية الدقة وتبقى عرضة لتوليد بيانات زائفة (الهلوسة) أو خوارزميات غير منضبطة. تشير التقييمات إلى أن هذه القيود تتقلص بنويًا عند تفعيل معمارية «الوكلاء المتعددين» (Multi-Agent Systems). يعتمد هذا النمط التنظيمي على توزيع المهام التشغيلية بين مثيلات متوازية؛ حيث يتولى نموذج بناء الأكواد، ويُخصص نموذج آخر للتدقيق المتقاطع، بينما ينفذ الثالث عمليات التنقيب التقني. تُظهر التقديرات ارتفاعاً ملحوظاً في معدلات الكفاءة لتتراوح بين 85% و95% ضمن هذه البيئات حيث يقارب هذا الأداء إحصائياً متوسط الكفاءة

المتوقع من طواقم التطوير البشري التقليدية. وأن القيمة الاستراتيجية للوكلاء المستقلين لا تكمن في تقديم منتج هندسي صفري الأخطاء من المحاولة الأولى، بل قد تتركز الفاعلية الحقيقية في الضغط الجذري للأطر الزمنية لدورة التصحيح (Debugging Cycle). ففي سيناريوهات الإخفاق الميداني كفشل اختبارات الإطلاق، وقد يتطلب من الكادر البشري فترات زمنية ممتدة قد تصل إلى أشهر لتشخيص الخلل في خوارزميات التوجيه. من المحتمل أن ينجز الوكيل الذكي هذه المهمة ذاتها، عبر توليد الفرضيات التشخيصية وإعادة كتابة الأكواد، خلال بضع ساعات. هذا التسريع في دورة «التجربة والخطأ» يضاعف وتيرة التراكم المعرفي ويرفع الكفاءة الحركية للمنظومة بسرعة فائقة. وترتبط استدامة هذه القدرات الهندسية بالانتقال التشغيلي نحو البيئات المعزولة محلياً. بمجرد توظيف الوكيل الذكي لبناء الخوارزميات الحسابية الأساسية وأنظمة الطيران الآلي، تُنقل هذه الحزم البرمجية للعمل على وحدات معالجة محلية (Offline) تمنح هذه الاستقلالية البنيوية التنظيمات المسلحة مناعة مباشرة ضد بروتوكولات الحظر التقني. يُتيح هذا المسار استمرار تشغيل أدوات المحاكاة، والاحتفاظ بنفس مستوى الدقة الهندسية، حتى في حال تعليق الحسابات أو قطع الوصول إلى واجهات الخدمات السحابية من قبل الشركات المزودة.

استغلال التكنولوجيا وآفاق التطوير المستقبلي

تشير الأنماط التشغيلية للفاعلين غير الحكوميين إلى استيعاب النماذج الخوارزمية كمعادل استراتيجي لتقليص الفجوة التكنولوجية حيال تفوق الجيوش النظامية. تتقاطع هذه الممارسة مع أربع مسارات تطويرية تُعيد تعريف ديناميكيات الصراع غير المتكافئ. فقد ارتبطت أنظمة التوجيه الدقيق وبرمجيات التحكم الجوي تاريخياً باقتصادات دفاع الدول الكبرى والبنى الأكاديمية المتخصصة. تُظهر المعطيات الميدانية أن وكلاء الذكاء الاصطناعي يوفرون مورداً هندسياً بديلاً عن المهندسين البشر. من المحتمل أن يُسهم هذا التدفق

المعرفي في تفكيك احتكار التصنيع العسكري المتقدم، مما يتيح لهذه المجموعات تجاوز الحواجز التقنية التي تفرضها بروتوكولات حظر التسلح وأنظمة الحصار الاقتصادي، وتحقيق قفزات مرحلية سريعة.

الاعتماد الأولي على واجهات الشركات التجارية لا يمثل سوى مرحلة انتقالية. تتجه مسارات التطوير الفعلي نحو تبني النماذج مفتوحة المصدر (Open-Source Mod-els). تضمن هذه المعمارية البرمجية إمكانية تحميل النماذج وتشغيلها محلياً ضمن وحدات معالجة معزولة تماماً عن شبكة الإنترنت (Air-gapped). يحدّد هذا التوجه آليات الرقابة الرقمية، ويتجاوز القيود الجغرافية وبروتوكولات الحظر التي تفرضها خوادم الشركات الغربية.

ينعكس هذا الاستقلال البرمجي مباشرة على هندسة المقذوفات المادية. يبرز توجه هندسي لتطوير منظومات ملاحية بديلة تتخلى عن إشارات التموضع العالمي (GPS). من خلال دمج معالجة البيانات الطرفية مع خوارزميات الرؤية الحاسوبية، تكتسب الطائرات المسيّرة استقلالية توجيهية للتعرف على التضاريس والأهداف. يوفر هذا التصميم حصانة تكتيكية صلبة ضد منصات الحرب الإلكترونية وأنظمة التشويش المصممة لقطع روابط الاتصال.

يمتد الأثر الهيكلي لهذه التقنيات ليصل إلى البنية التنظيمية لشبكات القيادة والسيطرة. تُقضي أتمّة مهام التحليل الاستخباري، ومعالجة الخرائط، وتنسيق الإمداد إلى قدرة تنظيمية على إدارة مسارح عمليات عسكرية واسعة وموزعة جغرافياً باستخدام طواقم بشرية مصغرة. الاعتماد على فرق عمل محدودة العدد يرفع بشكل ملحوظ من مستويات أمن العمليات (OPSEC). احتمالات الانكشاف الاستخباراتي تنقل طردياً مع تراجع الحاجة للتفاعلات البشرية الكثيفة والاتصالات الميدانية المعقدة.

السياسات الممكنة وآثارها على المستقبل

تفرض تطبيقات الذكاء الاصطناعي في قطاع التسليح إعادة هيكلة جذرية للتوازنات الجيوسياسية. تشير المعطيات إلى أن قدرة الفاعلين غير الحكوميين على تصميم خوارزميات التوجيه الدقيق وإدارة أسراب المسيّرات بتكلفة اقتصادية منخفضة تُؤسس لتآكل تدريجي في الفجوة التكنولوجية التي احتكرتها الدول الكبرى. يقترح هذا التحول تراجعاً محتملاً في فاعلية بروتوكولات الردع العسكري التقليدي. تتحول طبيعة الصراع غير المتماثل من الاعتماد على التكتيكات الفيزيائية الميدانية إلى نموذج هجين يقارب «حرب العصابات الرقمية». من المحتمل أن توفر وكلاء الذكاء الاصطناعي الهجومية للتنظيمات المسلحة مرونة تشغيلية وديناميكية تطوير تضاهي سرعة الاستجابة لدى المؤسسات الدفاعية النظامية.

صُممت معاهدات الحد من الانتشار ومراقبة التسليح المعيارية لتتبع حركة الكيانات الفيزيائية وسلاسل الإمداد المادية. بانتقال مركزية النقل العسكري نحو «المعرفة الخوارزمية» والبرمجيات التوجيهية، تبرز فجوة بنيوية في أطر الضبط الدولية. تقاوم الأصول الرقمية غير المادية محاولات التعقب الرقابي أو الحظر عبر المنافذ الحدودية، مما قد يُفقد آليات التفتيش الدولي فعاليتها المباشرة في تحجيم انتشار هذه القدرات.

تُظهر المؤشرات الاستراتيجية اتجاهاً متزايداً لدى القوى العظمى نحو تفعيل سياسات حمائية صارمة، تشمل تقييد حركة سلاسل توريد معالجات الذكاء الاصطناعي المتقدمة وتحجيم انتشار النماذج مفتوحة المصدر تحت مظلة الأمن القومي. يُرجح أن يُحفز هذا الضغط الاستراتيجي القوى المقابلة—تحديداً ضمن النطاقيين الروسي والصيني وامتداداتها الإقليمية—لتسريع بناء بيئات خوارزمية سيادية ومستقلة كلياً عن المعمارية التقنية الغربية. يُنذر هذا المسار بانسطار البنية التكنولوجية العالمية إلى معسكرين معزولين. قد تبرز هذه الثنائية الاستقطابية أسواقاً موازية غير نظامية (سوق سوداء رقمية) تنشط في تداول

النماذج التوليدية المهيأة للاستخدامات العسكرية البحتة.

التداعيات الاستراتيجية والأمنية

يُشير انهيار حاجز الكفاءة التقنية إلى تحولات عميقة في البنية الأمنية وقواعد الإسناد الاستخباراتي (Attribution). تقترح المعطيات الميدانية أن الاعتماد المتزايد على معمارية «الفريق الهجين» يُربك آليات الرصد التقليدية؛ حيث تتداخل الأنشطة البشرية مع المخرجات الآلية بطرق تعقد عمليات التتبع وتحديد المسؤولية بدقة وتظل القدرات الهندسية الجوهرية فاعلة ومستدامة حتى بعد تفعيل سياسات الحظر الرقمي. حيث تُفرز هذه التطورات إشكالية هيكلية تتعلق بالتناقض بين ديناميكيات الانفتاح التجاري للشركات التكنولوجية ومحددات الأمن القومي.

تتجاوز التقييمات الأمنية الحالية لمخاطر الذكاء الاصطناعي حدود التوظيف العسكري التكتيكي، لتشمل النمذجة الافتراضية لسيناريوهات فقدان السيطرة البنيوية. تشير التحليلات الصادرة في أيلول/سبتمبر 2026، ومن ضمنها طروحات الباحث السابق في «An-thropic» جيكوب كوكسون، إلى احتمالية وصول النماذج الفائقة لمستويات تعقيد قد تُقضي إلى عواقب كارثية. من المحتمل أن تتضمن هذه السيناريوهات اكتساب الخوارزميات قدرات تحكم غير مقيدة في البنى المادية، أو تنفيذ عمليات استتساخ ذاتي متوازية عبر الشبكات، مما يُعطل فاعلية بروتوكولات الإيقاف القسري (Kill switches). تبرز في هذا الإطار فرضيات تسخير هذه النماذج لهندسة عوامل بيولوجية كأداة محتملة لإحداث خسائر ديموغرافية حادة.

تتقاطع هذه الفرضيات مع تقديرات احتمالية طرحتها مرجعيات علمية وتكنولوجية. يُقدر «جيفري هينتون» احتمالية وقوع كارثة وجودية بنسبة تتراوح بين 10% إلى 20%، مرجحاً

ذلك لنشوء أهداف «حفظ الذات» (Self-preservation) كخصائص منبثقة (Emergent properties) داخل النماذج. تتماهى هذه النمذجة مع تقديرات «إيلون ماسك» التي تضع احتمالات الخطر عند عتبة الـ 20%. هذه الأرقام لا تعكس حتمية قاطعة. إنها تؤثر على تصاعد القلق المنهجي داخل مجتمعات التطوير.

يمتد هذا التحوّط ليشمل الأطر المؤسسية الدولية؛ إذ تعكس تصريحات المفوض السامي للأمم المتحدة لحقوق الإنسان، فولكر تورك، تنامي الاهتمام بالمخاطر الوجودية المحتملة المرتبطة بالذكاء الاصطناعي المتقدم. ويستند هذا التقييم إلى معطيات تشغيلية سُجّلت خلال مراحل النمذجة والاختبار، أظهرت فيها بعض الأنظمة سلوكيات تكيفية تمثّلت في محاولة تجاوز بيانات العزل المختبرية (Sandbox Evasion)، أو اتباع مسارات تنفيذية قد تعيق قدرة المشغلين على إيقاف تشغيلها.

الخاتمة

تشير المعطيات التحليلية المستمدة من التقييمات الأمنية لعام 2026 إلى تحوّل هيكلي في طبيعة التهديدات الاستراتيجية. فلم تعد النماذج اللغوية التوليدية مجرد أدوات مساندة، بل تشير المؤشرات الميدانية إلى احتمال تحولها إلى مكوّن هجومي مباشر يسهم في تقليص احتكار التكنولوجيا العسكرية. وتُظهر السجلات الموثّقة، الممتدة من المناطق الجغرافية في شمال اليمن إلى مقاطعة هونان الصينية، قدرة هذه الخوارزميات على تعزيز الكفاءة التشغيلية للخلايا التنظيمية المصغّرة، لتقترب مخرجاتها، من الناحية الإحصائية، من مستويات الأداء المتوقعة لدى الفرق الهندسية المتكاملة. ومن شأن هذا التحول أن يعيد تشكيل معادلات الردع الاستراتيجي التقليدية.

وفي المدى الزمني المنظور (2026-2028)، يُحتمل أن تتسارع وتيرة تبني النماذج مفتوحة المصدر القابلة للتوظيف في المجالات العسكرية. ويتزامن هذا المسار مع تطور معمارية «الفريق الهجين» متعدد الوكلاء، وهي بنية تنظيمية تتيح إدارة مشاريع تسليحية متفاوتة التعقيد بالاعتماد على عدد محدود من الكوادر البشرية. وتشير اتجاهات التطوير إلى احتمال تصاعد محاولات التحايل المنهجي على بروتوكولات الأمان الداخلية للشركات المطوّرة. كما قد تكتسب هذه القدرات قدراً من الاستقلالية التشغيلية من خلال إمكان تشغيلها ضمن بيانات محلية معزولة عن الشبكة (Offline)، بما يسمح باستمرار عمل بعض الأنظمة والبرمجيات حتى بعد تطبيق آليات الحظر الرقمي. وفي ظل هذه اللامركزية، قد تتراجع تدريجياً فاعلية سياسات الاحتواء التقني الشامل القائمة حصرياً على تقييد الوصول إلى الخدمات السحابية.

تفرض هذه المعطيات تجاوز المقاربات الأمنية القائمة حصراً على أدوات الرصد والتصفية الشبكية. وتتطلب إدارة هذا المستوى من التهديدات غير المتماثلة إعادة هندسة شاملة لأطر الأمن الجماعي. ويقترح هذا التحليل ضرورة وضع معايير دولية صارمة قادرة على استيعاب ديناميكيات التسلح الخوارزمي المتسارعة. وتمثل الاستجابة التنظيمية الاستباقية مساراً حتمياً لتأطير هذه المنظومات قبل بلوغها مستويات متقدمة من الاستقلالية التشغيلية عن الرقابة والتوجيه البشريين.

ويجب الاستثمار في التعليم التقني المتقدم، وتأسيس حاضنات وطنية لتطوير برمجيات الذكاء الاصطناعي مفتوحة المصدر (Open Source)، بما يتيح للكوادر المحلية بناء نماذج سيادية وأمنة، تعمل في بيئات معزولة، بعيداً عن الاعتماد المطلق على البنى السحابية الأجنبية المعرضة للحظر. كما ينبغي توجيه القدرات الهندسية نحو توظيف خوارزميات الرؤية الحاسوبية ومعالجة البيانات الطرفية في مجالات حماية الحدود، ومراقبة البنية التحتية، وإدارة الطوارئ، بكفاءة تواكب سرعة المعالجة الآلية الحديثة. ويُضاف إلى ذلك ضرورة تعزيز التدابير الأمنية وتطبيق معايير صارمة لحماية شبكات الطاقة والاتصالات والقطاع المالي والمصرفي، بما يسهم في التصدي لهجمات البرمجيات الخبيثة وشبكات التجسس المتقدمة المؤتمتة بالذكاء الاصطناعي.



لِدَوْلِيَّةٍ فَاعِلَةٍ وَمَجْتَمَعٍ مُشَارِكٍ

www.bayancenter.org

info@bayancenter.org
