

NAZA: عندما يصبح الفلسطيني Score — الذكاء الاصطناعي ونسبة الخطأ وثمان الاستهداف في غزة



عندما نشاهد منزلاً في غزة بعد قصفه، نرى عادة نهاية العملية: بناية مدمرة، سيارات إسعاف، جرحى، أشلاء، عائلات تبحث تحت الركام، وأسماء جديدة تُضاف إلى قوائم القتلى.

لكن ماذا حدث قبل سقوط القنبلة؟

من حدد الشخص المستهدف؟ من جمع معلوماته؟ هل جرى تحليل هاتفه واتصالاته وتحركاته؟ هل منحته خوارزمية ما درجة أو Score؟ كم كانت نسبة احتمال أن يكون هذا التصنيف خاطئاً؟ كم مدنياً كان متوقعاً وجودهم داخل المنزل؟ ومن اتخذ القرار في النهاية، وفي كم ثانية؟

هذه هي المنطقة التي يدخل إليها الفيلم الوثائقي NAZA.

وفي هذا المقال سأحاول أن أشرح، بلغة يفهمها غير المتخصصين، كيف تعمل هذه المنظومة من الناحية التقنية، ولماذا لا تكون نسبة الخطأ فيها مجرد رقم في تقرير أداء، بل مسألة تمس جوهر القانون الدولي الإنساني.

ما هو فيلم NAZA؟

NAZA فيلم وثائقي من كتابة وإخراج يوفال أبراهام وراحيل شور، مخرجي فيلم «لا أرض أخرى» الحائز على الأوسكار. عُرض للمرة الأولى في مهرجان البندقية السينمائي في سبتمبر/أيلول 2026 وحصل على جائزة لجنة التحكيم الخاصة، وهو من إنتاج JW Films و The Guardian.

بين عامي 2023 و2025، أجرى المخرجان مقابلات مع 24 جنديًا وضابطًا عملوا داخل المؤسسة العسكرية والاستخباراتية الإسرائيلية، مع إخفاء هويات عدد منهم لحمايتهم.

الفيلم لا يركز على لحظة سقوط القنبلة، بل على المنظومة التي تسبقها: المراقبة، الذكاء الاصطناعي، تحديد الأشخاص، إنتاج الأهداف، حساب الخسائر المدنية المتوقعة، والمراجعة البشرية قبل الضربات.

أما اسم **NAZA** نفسه، فتقول The Guardian إنه مصطلح عسكري استخباراتي يشير إلى عدد المدنيين المتوقع مقتلهم في ضربة جوية.

هذه الفكرة وحدها تختصر جانبًا كبيرًا من الفيلم: المدنيون ليسوا دائمًا رقمًا يُحصى بعد القصف. في بعض الحالات يكون عددهم المتوقع موجودًا قبل وقوعه.

ومن المقرر أن تتيح The Guardian الفيلم مجانًا على موقعها بعد انتهاء عرضه في صالات السينما أواخر نوفمبر/تشرين الثاني.

لا تتخيل روبوتًا يضغط زر القصف

عند سماع عبارة «الذكاء الاصطناعي في الحرب»، قد يتخيل البعض كمبيوترًا يقرر بمفرده إطلاق صاروخ.

ما تكشفه الشهادات والتحقيقات أكثر تعقيدًا. نحن أمام سلسلة من الأنظمة المترابطة، أو ما يسميه المتخصصون **Data Pipeline**، أي خط إنتاج تمر فيه المعلومات من مرحلة إلى أخرى:

بيانات ومراقبة

↓

ربط الأشخاص بالأجهزة

↓

تحليل الأنماط (Machine Learning)

↓

إعطاء درجة (Score)

↓
اقتراح أهداف

↓
تتبع الهدف

↓
تقدير المدنيين المتوقع وجودهم

↓
مراجعة بشرية

↓
ضربة

احفظ هذه الصورة في ذهنك، لأننا سنعود إليها. فالسؤال الأهم في هذا المقال ليس فقط: هل تخطئ إحدى هذه المراحل؟ بل: ماذا يحدث عندما تخطئ عدة مراحل متتالية؟

المرحلة الأولى: الإنسان كبيانات

أي نظام ذكاء اصطناعي يحتاج إلى بيانات. وفي أنظمة الاستخبارات قد تأتي البيانات من الهواتف، والمكالمات، والمواقع الجغرافية، والصور، والعلاقات بين الأشخاص، وأنماط الحركة، وقواعد البيانات السابقة.

داخل النظام، يمكن أن يبدو الإنسان هكذا:

Person
├─ Name
├─ Phone
├─ SIM
├─ Home
├─ Location
├─ Contacts
├─ Calls
├─ Family
└─ Movement History

ثم تُربط هذه الملفات ببعضها:

Person A — اتصل بـ — Person B



هذا ما يُعرف بتحليل الشبكات (Graph Analysis). في التطبيقات المدنية يُستخدم لكشف الاحتيال المصرفي مثلاً. وفي العمل الاستخباراتي يُستخدم لرسم خريطة علاقات بين الناس.

بالنسبة للمبرمج، هذه مجرد كيانات وعلاقات في قاعدة بيانات. لكن كل نقطة هنا إنسان حقيقي.

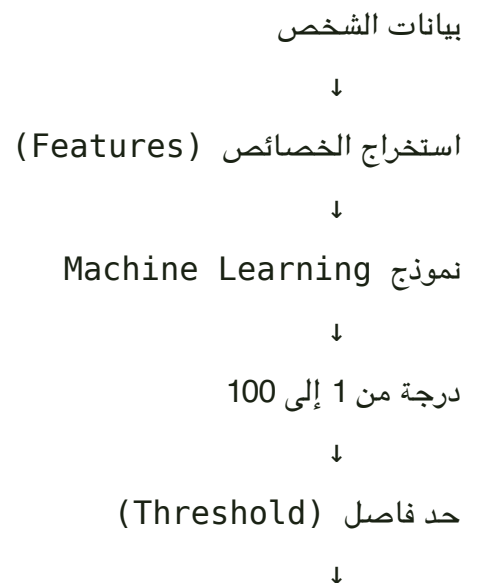
ومن هنا يبدأ السؤال الحساس: هل الاتصال بشخص ما يجعلك جزءاً من مجموعته؟ هل استخدام هاتف كان يخص شخصاً آخر يجعلك مشتبهاً به؟ في مجتمع صغير ومحاصر مثل غزة، حيث يعرف الجميع الجميع، كم شخصاً يمكن أن يكون «مرتبطاً» بأحدهم؟

Lavender: عندما يتحول الإنسان إلى Score

في أبريل/نيسان 2024، نشرت مجلة +972 وموقع Local Call تحقيقاً قال فيه ستة ضباط استخبارات إسرائيليين إن نظاماً يُعرف باسم **Lavender** استُخدم لتصنيف الفلسطينيين في غزة بحسب احتمال ارتباطهم بالجناح العسكري لحماس أو الجهاد الإسلامي.

بحسب التحقيق، حلل النظام معلومات عن أعداد كبيرة من سكان غزة، وأعطى كل شخص درجة من 1 إلى 100 تعكس، وفق النموذج، احتمال أن يكون عنصراً في إحدى التنظيمات.

ببساطة:



ما معنى Feature؟

ال-Feature هي أي معلومة يعتمد عليها النموذج في قراره. بحسب التحقيق، تحدثت المصادر عن خصائص مثل أنماط الاتصال، والعلاقة بأشخاص معروفين، وتغيير الهاتف أو العنوان.

النظام يتعلم من أشخاص صنّفهم الاستخبارات مسبقاً، ثم يبحث عن أشخاص آخرين «يشبهونهم».

لكن فكر في هذا: في حرب نزع فيها معظم سكان غزة مراراً، تغيير العنوان ليس سلوكاً مريباً، بل هو حياة الجميع. وتغيير الهاتف قد يعني فقط أن الهاتف القديم ضاع تحت الركام.

التشابه لا يعني التطابق. وهذه من أبسط حقائق Machine Learning.

نسبة الخطأ 10%: ماذا تعني فعلاً؟

بحسب المصادر التي تحدثت إلى +972 Local Call، اختبر الجيش عينة من مخرجات Lavender، ووجد أن التصنيف كان صحيحاً في نحو 90% من الحالات.

قد يبدو هذا رقماً جيداً. لكن يجب أن نفهم بالضبط ما الذي يقيسه.

هذه النسبة لا تعني أن النظام يخطئ مرة واحدة كلما فحص عشرة أشخاص عشوائيين من سكان غزة. هي تعني أنه من بين كل عشرة أشخاص وضعهم النظام على قائمة الأهداف، هناك واحد على الأقل لا ينتمي إلى الفئة التي يُفترض استهدافها.

في لغة المختصين يُسمى هذا **Precision**: من بين كل من قال النظام إنهم «أهداف»، كم منهم أهداف فعلاً؟

مثال من المختبر الطبي

تخيل أن مختبراً أخبرك أن نتيجة تحليلك إيجابية، وأن نتائجه الإيجابية خاطئة في 10% من الحالات.

هل تقبل أن تُجرى لك عملية جراحية خطيرة بناءً على هذه النتيجة وحدها، دون فحص ثانٍ؟

أي طبيب مسؤول سيطلب فحصاً تأكيدياً. هذا هو الحد الأدنى حين تكون العواقب غير قابلة للتراجع.

في منظومة الاستهداف، «العملية الجراحية» قبلية. والسؤال هو: هل كان هناك فحص ثانٍ حقيقي؟ سنعود لهذا بعد قليل.

الحساب بالأرقام الحقيقية

بحسب نفس التحقيق، وصل عدد الفلسطينيين الذين صَنَّفهم Lavender في إحدى مراحل الحرب كأهداف محتملة إلى نحو 37 ألف شخص.

لا نحتاج إلى أمثلة افتراضية. يكفي أن نطبق نسبة الخطأ التي اعترف بها الجيش نفسه على هذا الرقم:

$$37,000 \text{ شخص على قائمة الأهداف المحتملة} \\ \times 10\% \text{ نسبة خطأ (بحسب اختبار الجيش نفسه)}$$

≈ 3,700 إنسان

أي أن نحو 3,700 شخص ربما وُضعوا على قائمة الأهداف وهم ليسوا من يظنهم النظام.

هذا لا يعني بالضرورة أن 3,700 شخص قُتلوا بسبب خطأ Lavender، فلا نملك بيانات عن عدد من استُهدف فعلاً من القائمة. لكنه يعني أن المنظومة عملت وهي تعرف مسبقاً أن آلاف الأبرياء على الأرجح موجودون على قائمتها.

والخطأ الحقيقي قد يكون أكبر من 10%

هنا نقطة تقنية دقيقة لكنها مهمة جداً.

نسبة الـ 90% «صحيحة» مقارنة بماذا؟

أي نموذج Machine Learning يُقاس مقابل ما يسمى **Ground Truth**، أي «الحقيقة» التي نعتبرها صحيحة. وأحد المصادر في تحقيق +972 قال إن تعريف «عنصر حماس» المستخدم في بعض بيانات تدريب Lavender كان واسعاً، وإن بيانات أشخاص مثل عاملين في الدفاع المدني دخلت ضمن عملية التدريب.

إذا صح ذلك، فالنظام قد يكون «دقيقاً» بنسبة 90% وفق تعريف فضفاض أصلاً. أي أنه قد يصيب في تصنيف شخص كـ «مرتبط بـ حماس» وفق هذا التعريف، بينما هذا الشخص ليس مقاتلاً بالمعنى القانوني.

تعريف واسع وغير دقيق للهدف

↓

بيانات تدريب منحازة

↓

نموذج يتعلم الانحياز

↓

تصنيف خاطئ على نطاق واسع



لكن الاختبار يقول: 90% دقة!

هذا ما يسميه المبرمجون: **Garbage In, Garbage Out**. إذا كانت البيانات الداخلة خاطئة، لن يحولها الذكاء الاصطناعي إلى حقيقة. بل سيجعل الخطأ أسرع، وأكثر اتساقاً، وعلى نطاق أكبر.

بمعنى آخر: نسبة الخطأ بالمعيار القانوني، أي «هل هذا الشخص مقاتل فعلاً؟»، قد تكون أعلى من 10%، لأن المسطرة التي قيس بها النظام نفسها معوجة.

من يحدد الحد الفاصل؟

إذا أعطى النظام كل إنسان درجة من 1 إلى 100، فلا بد من حد يقرر متى يتحول الشخص إلى هدف:

Score $\geq$ 90 → هدف

Score < 90 → تجاهل

لكن لماذا 90؟ لماذا ليس 95 أو 99 أو 70؟

هذه ليست حقيقة علمية، بل قرار يتخذه بشر. وكل تخفيض للحد يعني قائمة أطول، ونسبة خطأ أعلى.

بحسب أحد المصادر في تحقيق +972، كان تخفيض الحد يؤدي إلى إنتاج عدد أكبر من الأهداف، وأشار إلى وجود ضغط لإنتاج مزيد منها.

قرار تغيير رقم واحد في إعدادات النظام يمكن أن يضيف آلاف البشر إلى قائمة المشتبه بهم.

حين يصبح الإنسان هو «عنق الزجاجة»

في أي نظام تقني كبير، نبحث عن **Bottleneck**، أو عنق الزجاجة: الجزء الذي يبطئ كل شيء.

في منظومة إنتاج الأهداف، كان الإنسان نفسه هو عنق الزجاجة.

التحقيق أشار إلى كتاب بعنوان *The Human-Machine Team*، نُشر باسم مستعار (Brigadier General Y.S). ونُسب إلى قائد الوحدة 8200 الإسرائيلية، يتحدث عن الحاجة إلى الذكاء الاصطناعي لإنتاج أعداد كبيرة من الأهداف، لأن البشر لا يستطيعون معالجة هذا الحجم من البيانات.

الهدف من الأتمتة عادة هو إزالة عنق الزجاجة.

لكن إذا كان عنق الزجاجة هو الإنسان الذي يراجع الأدلة قبل قتل إنسان آخر، فهل إزالته تحسين للنظام؟ أم إزالة لآخر صمام أمان فيه؟

Human in the Loop ... لمدة 20 ثانية

أحد أهم الردود على انتقاد استخدام الذكاء الاصطناعي عسكرياً هو أن الإنسان يبقى صاحب القرار النهائي. وهذا يسمى **Human in the Loop**.

تذكر مثال المختبر: الفحص الثاني هو ما يحمي المريض من الـ 10% الخطأ. فكيف كان يبدو «الفحص الثاني» هنا؟

وفق أحد المصادر في تحقيق +972، كان يخصص قرابة 20 ثانية لمراجعة بعض الأهداف، وقال إن دوره اقتصر في حالات على التأكد من أن الشخص الذي اختاره النظام رجل. ووصف نفسه بأنه لم يكن أكثر من ختم موافقة.

جيش الاحتلال الصهيوني رفض هذه الرواية، وقال إن المحلل مطالب بإجراء مراجعة مستقلة والتحقق من شرعية الهدف وفق التعليمات والقانون الدولي.

لكن إذا صحت الشهادة، فالتحقق من جنس الشخص لا يكشف شيئاً من الـ 10% الخطأ. الرجل الذي صُنّف خطأً يبقى رجلاً. هذا ليس فحصاً ثانياً، بل تأكيد شكلي للفحص الأول.

عندما نصدّق الكمبيوتر أكثر من اللازم

هناك ظاهرة معروفة في علم الأنظمة تسمى **Automation Bias**: ميل الإنسان إلى الوثوق بنتيجة الآلة، خاصة عندما تبدو دقيقة وعلمية.

تخيل أن هذه الشاشة ظهرت أمام محلل لديه 20 ثانية (مثال توضيحي من إعداد الكاتب، وليس صورة من النظام الحقيقي):

Person ID: 182736

Affiliation Score: 94%

Location: Building 428

Civilian Estimate: 11

Status: Target Candidate

كل شيء يبدو مرتباً. هناك أرقام، ونسبة مئوية، وموقع.

لكن الـ 94% ليست حقيقة. إنها ناتج نموذج بُني على بيانات وتعريفات وحدود وضعها بشر. وفي 20 ثانية، لا أحد

يستطيع أن يسأل: من أين جاءت هذه الـ94%؟

Where's Daddy?: عندما يعود الإنسان إلى بيته

بعد تحديد الهدف، تأتي مشكلة أخرى: أين هو الآن؟

بحسب التحقيق، استُخدم نظام سُمّي **Where's Daddy** لمتابعة الأشخاص المحددين وإرسال إشارة عندما يدخلون منازل عائلاتهم.

رقم الهدف

↓

تتبع الموقع

↓

الهدف دخل منزله

↓

إشارة (Event)

↓

بدء إجراءات الضربة

كل مطور تطبيقات يعرف هذا النمط، ويسمى **Event-Driven Architecture**: التطبيق ينتظر حدثاً معيناً، وعندما يقع ينفذ إجراءً.

لكن الحدث هنا: رجل عاد إلى بيته. والإجراء قد يكون ضربة جوية.

بحسب المصادر، كان استهداف الأشخاص في منازلهم أسهل استخباراتياً. البيت، المكان الذي يفترض أن يكون الأكثر أماناً ومدنية، يصبح الموقع الأسهل للتحديد. وعندما يدخل الرجل بيته، يكون حوله زوجته وأطفاله ووالداه وجيرانه.

NAZA: كم مدنياً يُتوقع أن يموت؟

هذه ربما أهم نقطة في الفيلم كله.

النظام لا يسأل فقط: هل يوجد هدف هنا؟ بل يسأل أيضاً: كم مدنياً سيُقتل إذا قصفنا هذا المكان؟

الهدف + الموقع + المبنى + السلاح + عدد السكان المقدّر

↓

عدد المدنيين المتوقع مقتلهم

↓

القرار

في قاعدة البيانات قد يظهر السطر هكذا:

Expected civilian casualties: 17

بالنسبة للنظام، 17 عدد صحيح (Integer). بالنسبة للتقرير، إحصائية. بالنسبة لعائلة فلسطينية: أب وأم وأطفال وأشقاء وأقارب.

وبحسب مصدرين في تحقيق +972، سمح الجيش خلال الأسابيع الأولى من الحرب، في بعض عمليات استهداف العناصر منخفضة الرتبة، بقتل ما يصل إلى 15 أو 20 مدنيًا مقابل الشخص المستهدف. وقالت المصادر إن الحد كان أعلى بكثير عند استهداف القيادات، ووصل في حالات إلى أكثر من مئة مدني.

هذه شهادات مصادر، ولا تعني أن رقمًا واحدًا طبق على كل ضربة، وقد رفض الجيش الإسرائيلي عددًا من توصيفات التحقيق. لكن إن صحت، فنحن لا نتحدث عن مدنيين ماتوا في حدث مفاجئ، بل عن موت محسوب مسبقًا.

الأخطاء لا تحدث مرة واحدة... بل تتراكم

حتى الآن تحدثنا عن خطأ Lavender وحده. لكن عدّ إلى خط الإنتاج في بداية المقال. كل مرحلة فيه لها هامش خطأ خاص بها. وبحسب نفس التحقيق، تحدثت المصادر عن أخطاء في مراحل أخرى:

- خطأ في الهوية: النظام يتتبع الهاتف وليس الإنسان. وفي غزة يتبادل الناس الهواتف، وقد ينتقل هاتف إلى فرد آخر من العائلة بعد مقتل صاحبه.
- خطأ في التوقيت: في بعض الحالات مر وقت بين إشارة دخول المنزل وتنفيذ الضربة، فكان الشخص المستهدف قد غادر، وقُتلت عائلته وحدها.
- خطأ في تقدير المدنيين: تقدير عدد من في المبنى اعتمد على بيانات تقريبية، في مدينة نزع فيها الناس إلى بيوت أقاربهم وتضاعف عدد سكان كل بيت.
- خطأ في السلاح: بحسب المصادر، كانت الأهداف منخفضة الرتبة تُضرب غالبًا بقنابل غير موجهة لأنها أرخص، وهي أقل دقة وأوسع أثرًا.

والمشكلة أن هذه الأخطاء لا تُجمع، بل تُضرب ببعضها.

لنأخذ مثالاً توضيحيًا (الأرقام هنا افتراضية لشرح الفكرة، باستثناء نسبة Lavender): لو كانت كل مرحلة من أربع مراحل صحيحة بنسبة 90%:

$$\begin{aligned} & \text{التصنيف} \times \text{الهوية} \times \text{التوقيت} \times \text{تقدير المدنيين} \\ & 0.9 \times 0.9 \times 0.9 \times 0.9 \\ & \approx 0.66 \end{aligned}$$

أي أن سلسلة من أربع مراحل «جيدة»، كل واحدة منها صحيحة في تسع حالات من عشر، قد تنتج قراراً صحيحاً بالكامل في ثلثي الحالات فقط.

هذه قاعدة أساسية في هندسة الأنظمة: موثوقية السلسلة أقل دائماً من موثوقية أضعف حلقاتها.

Scale: حين يتضاعف الخطأ

البشر يخطئون أيضاً. هذا صحيح.

لكن المحلل البشري قد يراجع عشرات الحالات. أما النظام الآلي فيعالج عشرات الآلاف.

خطأ بشري × عشرات القرارات

مقابل

خطأ آلي × 37,000 تصنيف

الأمثلة لا تزيد السرعة فقط. إنها تضخم كل خطأ موجود في التصميم أو البيانات أو السياسة، وتحوله من حادثة فردية إلى نمط منهجي.

من الخطأ التقني إلى جريمة الحرب

هنا يتوقف النقاش عن كونه تقنياً فقط.

القانون الدولي الإنساني يقوم على ثلاثة مبادئ أساسية: التمييز بين المدنيين والمقاتلين، والتناسب بين الضرر المدني والميزة العسكرية، والاحتياط لتقليل الضرر على المدنيين.

الإحتلال الصهيوني ليست طرفاً في البروتوكول الإضافي الأول لاتفاقيات جنيف الذي فصل هذه المبادئ، لكن هذه المبادئ الأساسية تُعتبر جزءاً من القانون الدولي العرفي الملزم لكل الدول. ولنر كيف يصطدم كل رقم تقني بها:

1. عند الشك، هو مدني

ينص البروتوكول الإضافي الأول (المادة 50) على أنه إذا ثار شك حول ما إذا كان الشخص مدنيًا، فيُعتبر مدنيًا.

نظام يعمل بهامش خطأ معروف يبلغ 10% على الأقل، ويُراجَع في 20 ثانية، يقلب هذه القاعدة رأسًا على عقب: الشخص يُعامل كهدف إلى أن يثبت العكس، ولا أحد لديه الوقت ليثبت العكس.

2. واجب التحقق

يُلزم البروتوكول نفسه (المادة 57) من يخطط لهجوم بأن يفعل كل ما هو ممكن عمليًا للتحقق من أن الهدف عسكري.

السؤال بسيط: هل التأكد من أن الشخص رجل، خلال 20 ثانية، هو «كل ما هو ممكن»؟ خاصة حين يعرف الجيش نفسه أن واحدًا من كل عشرة على القائمة خطأ؟

3. التناسب ينهار حين يكون الهدف خطأ

هذه ربما أهم نقطة تقنية-قانونية.

مبدأ التناسب معادلة: الضرر المدني المتوقع في كفة، والميزة العسكرية في الكفة الأخرى. عندما يُسمح بقتل 15 أو 20 مدنيًا لاستهداف شخص واحد، يكون هذا الحساب مبنياً على افتراض أن هذا الشخص هدف عسكري فعلاً.

لكن ماذا لو كان من الـ 10% الخطأ؟

الضرر المدني المتوقع: 20 إنساناً
الميزة العسكرية: 0 (الشخص ليس مقاتلاً)

النتيجة: 21 مدنيًا قُتلوا مقابل لا شيء

في هذه الحالة، كل من قُتل في الضربة مدنيون، بمن فيهم الشخص «المستهدف» نفسه. معادلة التناسب لا تصبح مختلة فقط، بل تنهار كلياً.

وعندما تعمل المنظومة وهي تعرف أن هذا سيحدث في نحو واحدة من كل عشر حالات على الأقل، فهذا ليس احتمالاً نظرياً، بل نتيجة إحصائية مؤكدة على نطاق عشرات الآلاف.

4. الخطأ المعروف مسبقاً ليس حادثاً

في الحرب، لا يُعتبر كل خطأ جريمة. الجندي الذي يخطئ في ظروف لم يكن يستطيع توقعها يختلف عن الجهة التي

تعرف مسبقاً حجم الخطأ وتقبله.

نظام روما الأساسي للمحكمة الجنائية الدولية يعتبر من جرائم الحرب تعمّد توجيه هجمات ضد السكان المدنيين، وشنّ هجوم مع العلم بأنه سيسبب خسائر مدنية مفرطة بشكل واضح مقارنة بالميزة العسكرية.

والكلمة المفتاحية هنا: العلم.

عندما تختبر جهة ما نظامها، وتعرف أنه يخطئ في واحد من كل عشرة، ثم تقرر استخدامه على عشرات الآلاف، مع مراجعة بشرية شكلية، وهامش مسبق يسمح بقتل عشرات المدنيين في كل ضربة، واستهداف متعمد للبيوت ليلاً حيث تكون العائلات، فإن موت الأبرياء لم يعد «خطأ». أصبح نتيجة معروفة ومقبولة سلفاً، مبنية داخل تصميم النظام نفسه.

بين الاتهام والحكم

من المهم أن نكون دقيقين: إثبات أن ضربة بعينها جريمة حرب يحتاج إلى تحقيق في ظروفها وأدلتها وامتخذ القرار فيها.

لكن المشكلة التي يكشفها NAZA والتحقيقات المحيطة به ليست في ضربة واحدة. إنها في المنظومة نفسها، كما تصفها شهادات أشخاص عملوا داخلها: منظومة صُممت بطريقة تجعل قتل مدنيين بالخطأ نتيجة متوقعة ومحسوبة ومقبولة. وهذا بالضبط ما يجعلها موضع مساءلة قانونية، لا مجرد نقاش أخلاقي.

وهذه المسألة ليست افتراضية. في 21 نوفمبر/تشرين الثاني 2024، أصدرت المحكمة الجنائية الدولية مذكرتي توقيف بحق رئيس الوزراء الإسرائيلي بنيامين نتنياهو ووزير الدفاع السابق يوآف غالانت.

وقالت الدائرة التمهيدية إن هناك أسباباً معقولة للاعتقاد بمسؤوليتهما عن جرائم حرب وجرائم ضد الإنسانية، منها جريمة الحرب المتمثلة في التجويع كأسلوب من أساليب الحرب، كما وجدت أسباباً معقولة للاعتقاد بمسؤوليتهما كمسؤولين مدنيين عن جريمة الحرب المتمثلة في تعمّد توجيه هجوم ضد السكان المدنيين.

مذكرة التوقيف ليست حكماً بالإدانة. لكنها تعني أن الأسئلة التي يطرحها الفيلم موجودة بالفعل داخل مسار قضائي دولي.

الفالسطيني يرى النتيجة... والنظام يرى البيانات

الفالسطينيون، صحفيين وأطباء وعمال إنقاذ وعائلات ضحايا، تحدثوا منذ اليوم الأول عن حجم الخسائر المدنية. لكن عندما يشرح شخص من داخل المؤسسة العسكرية آلية إنتاج الهدف، تحصل المعلومة على اهتمام دولي مختلف. وهذا بحد ذاته يطرح سؤالاً عن من تُمنح شهادته وزناً أكبر.

ومع ذلك، تبقى هذه الشهادات مهمة لأنها تكشف ما لا تستطيع الضحية رؤيته: الجانب الآخر من الشاشة.

الفلسطيني يرى:	النظام يرى:
بيتاً	Entity
عائلة	Device
طفلاً	Location
جاراً	Feature
اسماً	Score
قصة	Target
	Collateral Damage Estimate

الرواية الفلسطينية تقول: هذا المنزل كان فيه عائلة.

وNAZA يدفعنا لنسأل: ماذا كان يظهر على شاشة من قرر ضربه؟ هل ظهر عدد المدنيين؟ هل ظهرت نسبة ثقة؟ هل كان الشخص المستهدف في البيت أصلاً؟ وكم ثانية استغرقت المراجعة؟

الخلاصة: خلف ال-Score يوجد إنسان

هذه كلمات مألوفة لكل من يعمل في البرمجة:

Data → Feature → Model → Score → Threshold → Target → Trigger → Approve

كلها تبدو تقنية ومحيدة. لكن ضعها داخل منظومة عسكرية وستتغير معانيها:

ال-Data قد تكون حياة فلسطيني كاملة جُمعت عبر المراقبة.

وال-Feature قد تكون مكالمات هاتفية، أو هاتفاً جديداً بعد أن ضاع القديم تحت الركام.

وال-Score قد يحدد مدى الاشتباه به.

وال-Threshold قد يقرر إن كان سينتقل إلى قائمة الأهداف.

وال-Tracking قد يعرف متى عاد إلى بيته.

وال-NAZA قد يمثل عدد أفراد عائلته المتوقع مقتلهم معه.

ثم يأتي إنسان في نهاية السلسلة، لديه 20 ثانية، ليضغط Approve.

وكل ذلك يحدث بينما يعرف أصحاب النظام أن واحداً على الأقل من كل عشرة على القائمة ليس من يظنونه.

في البرمجة، عندما نكتشف خطأ نصدر Patch. نعيد تدريب النموذج. نغير ال-Threshold. نعمل Rollback إلى

الإصدار السابق.

لكن عندما يكون ناتج ال-Output قنبلة سقطت على منزل، لا يوجد Rollback.

ولا يبقى ال-False Positive رقماً في تقرير أداء.

يصبح اسماً على شاهد قبر.