MATTHEW
HALABA.

ANALYTICAL REPORT
AS OF 25.11.2026

RUSSIA–UKRAINE WAR · OPEN SOURCES

AI on the Battlefield in Ukraine

What open data confirms, and what it cannot show

Analytical report. As of 25 November 2026. Figures and quoted claims were re-checked against the sources on that day. Scope: open sources on the Russia–Ukraine war, with no instructions for building weapons or carrying out attacks.

Matthew Halaba · halaba.online · matthew@halaba.online

Edition 25.11.2026

CONTENTS

In this report

1	Conclusions	3
2	How the evidence was assessed	6
3	Verification of key data	7
4	Map of applications	11
5	Analysis	13
5.1	The sensor-to-shooter loop and the denominator problem	13
5.2	Autonomy: function by function	15
5.3	Electronic warfare as a constraint	17
5.4	Data, labelling and model drift	18
5.5	Legal and model-safety risks	19
5.6	Influence operations: use is not effect	21
5.7	Exercises versus war	21
6	Marketing versus evidence	22
7	Panel claims tested against the evidence	23
8	Recommendations for Poland and NATO	25
9	Ten questions without a reliable public answer	27
10	Scope of verification and limitations	28
—	Sources	29

TITLE	AI on the Battlefield in Ukraine. What open data confirms, and what it cannot show
AS OF	25 November 2026 (figures and claims re-checked on that day)
SCOPE	Open sources on the Russia–Ukraine war, with no instructions for building weapons or carrying out attacks
SOURCES	18 items; numbers in square brackets link to the source list
EDITION BY	Matthew Halaba · halaba.online

1 Conclusions

Main thesis

Open data confirm that AI is genuinely being used in Ukraine for image analysis, data fusion and terminal guidance of drones. They do not, however, make it possible to measure how much AI improves effectiveness compared with a well-organised system without it. The advantage comes from the whole ecosystem — data, communications, integration, people and the pace of iteration — not from a single algorithm.

FIGURES THAT RECUR IN THIS REPORT

~12k
equipment detections per week by Avengers (MoD claim), with no accuracy metrics
Ukrainian MoD, 23.09.2024 [1]

10–80%
spread of claimed FPV success rates, calculated against different denominators
CSIS, Jamestown [2][14]

3–8
drones: the typical size of a "swarm" in practice — a small group, not collective intelligence
Ifri, February 2026 [10]

70→6%
fall in the effectiveness of Excalibur shells at the peak of the 2023 offensive owing to EW
Ifri [10]

HOW TO READ THE CONCLUSIONS

The eight conclusions on the following pages separate two things that usually merge in public debate: the fact that an AI function **is being used**, and the claim that it **produces a measured effect**. Each point gives a confidence level and source numbers; the full assessment method is set out in Chapter 2, and the verification of individual figures in Chapter 3.

Confirmed use

Image analysis, data fusion, terminal guidance of drones towards a target designated by a human.

Unmeasured effect

A universal "AI effectiveness multiplier", mass swarms, fully autonomous target selection as the norm.

1 Conclusions

POINTS 1–4

1 **Image analysis is the best-documented application.** Avengers, running in the VEZHA module of the DELTA system, automatically detects equipment in video from drones and fixed cameras. On 23.09.2024 the Ukrainian Ministry of Defence reported a volume of about 12,000 detections a week, but no accuracy metrics. Confidence: high for use, low for effect [1][2]

2 **Claimed success rates are inconsistent and use different denominators.** CSIS reports an increase from about 10–20% to about 70–80% after autonomous navigation was added, based on interviews. The same report, however, gives 10–15% for ordinary FPV missions and 30–50% for experienced operators. Jamestown, citing the MoD, gives 80% in 2026 against 30–50% at the start of the war. Confidence in a universal multiplier: low [2][14]

3 **In the Ukrainian cases described, a human selects the target.** Reuters (29.11.2025) described a drone that kept its guidance after losing its link. NPR (29.10.2026) watched a test in which the pilot designated a car and the drone took over the flight. Both reports confirm the function, not a success statistic. [5] [6]

4 **Russia's V2U is the strongest indication of wider autonomy, but not proof.** CSIS (April 2026) assessed that Russia had probably used a fully autonomous system. The basis is mainly Ukrainian analysis of captured units lacking communications components, together with battlefield accounts. In March 2025 the same think tank was still writing that autonomy in the sense of the US definition did not exist on the battlefield in Ukraine, so the change of assessment reflects both new data and the definition. [2][3]

1 Conclusions

POINTS 5–8

5 **True swarms are not a demonstrated norm.** RUSI (16.07.2026) distinguishes one operator controlling many drones from a swarm with collective intelligence, and does not see the latter at significant scale in Ukraine. Ifri notes that swarms are in practice limited to small groups, usually 3–8 drones. [7][10]

6 **Electronic warfare and physics still set the limits.** Terminal autonomy and fibre-optic links change which link in the chain is vulnerable to jamming. They do not make the whole mission resilient. [5][10]

7 **General-purpose models have weak safeguards against use in targeting.** The Vals AI study (17.11.2026) found that no prompt was blocked by a classifier, and that a "this is a game" framing removed some refusals. The study does not, however, measure the effectiveness of deployed systems or compliance with the law of war. [8][9]

8 **The best-founded direction for Poland and NATO is data, integration, testing under defined conditions and sustaining capability — not buying an "AI" label.** This is an analytical conclusion drawn from Ifri and CSIS, not a result measured in war. [2][10]

In short

In this war AI is a tool that speeds up the work of people and of whole systems. The most evidence concerns image analysis and guidance in the final phase of flight; the least concerns independent target selection and swarms. Every percentage has to be read together with what exactly it measures.

2 How the evidence was assessed

Each claim was assessed separately, because one can be highly confident that a system is in use and have low confidence that it delivers the claimed success rate.

Confirmed	Partly confirmed	Unconfirmed	Unverifiable	Not verified
The evidence is sufficient to confirm a precisely defined function or event.	Credible indications, but no independent verification of scale, outcome or mechanism.	The sources checked do not contain sufficient evidence.	Assessment requires non-public operational data.	The source could not be opened for this version of the report. Still to be checked.

Status scale used in the tables of Chapters 3, 4 and 7. The colour of the left edge of the "Assessment" cell matches the status.

Status	Meaning
Confirmed	The evidence is sufficient to confirm a precisely defined function or event.
Partly confirmed	There are credible indications, but independent verification of the scale, outcome or mechanism is lacking.
Unconfirmed	The sources checked do not contain sufficient evidence.
Unverifiable	Assessment requires non-public operational data.
Not verified	The source could not be opened for this version of the report. The claim remains to be checked.

Main bias of the material: the predominance of Ukrainian and Western perspectives. Information on Russian systems comes largely from analysis of captured equipment and from Ukrainian accounts.

Independence rule: several publications repeating the same manufacturer or ministry statement are not several confirmations. The CSIS and Ifri reports are solid studies based on interviews and field research, but they are not an audit of all combat missions. [2][10]

3 Verification of key data

The table sets what the report previously claimed against what the source says after re-checking. Rows marked CORRECTION change earlier wording.

Claim	What the source says	Assessment
Avengers: about 12,000 detections a week	MoD, 23.09.2024: 12,000 items of equipment a week from automatic analysis of video from drones and fixed cameras. The integration runs in the VEZHA module of the DELTA system. No accuracy or error metrics. [1]	Confirmed as a claim. The figure measures detections, not kills.
Autonomy raises success from 10–20% to 70–80%	CSIS, March 2025: increase based on interviews (the report mentions dozens of conversations). In the same report ordinary FPV missions are 10–15% and experienced operators 30–50%, so the data are not fully consistent. [2]	Confirmed as a claim. Unconfirmed as a front-wide average.
The Lyptsi operation proves autonomy	CSIS describes it as an uncrewed operation in which the systems were nevertheless operated manually and remotely. [2]	Confirmed. The inference of autonomy: wrong.
VEZHA replaces analysts' work	According to CSIS, people still do most of the analysis in VEZHA. [2]	Confirmed. Replacement of analysts: unconfirmed.
Reuters: drone kept its guidance after losing its link	Reuters, 29.11.2025: pilot "Mex" of the 58th Separate Rifle Brigade, a strike from about 20 km on what was probably a tank. The system locks onto the target image and flies on after contact is lost. The industry admits the software needs fixes and that effectiveness depends on conditions. Ukraine states that the decision to strike is made by a human. [5]	Confirmed for the application. No mission statistics.

3 Verification of key data

CONT. · ROWS 6–9

Claim	What the source says	Assessment
NPR: The Fourth Law test	NPR, 29.10.2026: the pilot selected the target with the stick, the drone signalled a lock, the pilot released the stick, and the drone reached a moving car. The test drone carried no payload. Dozens of units use the Lupynis-10-TFL-1. The slogan "10,000 times more effective" is the company founder's remark about asymmetric advantage, with no metric. [6]	Confirmed for the test. Multiplier: unconfirmed.
V2U is fully autonomous	CSIS, April 2026: "probably". Captured units from summer 2025 had LTE modems, while those from October and November 2025 no longer did. The on-board computer is a Jetson Orin. The author describes group coordination as inference and the design as primitive. Sources largely Ukrainian. [3]	Partly confirmed. Medium confidence.
Swarms in Ukraine	RUSI, Robert Tollast, 16.07.2026: swarms have been demonstrated in trials but not used at significant scale. Operation Spiderweb, using 117 partly autonomous drones, is cited as an example, but not as a swarm. [7]	Confirmed. Mass swarms: unconfirmed.
Vals AI: general models and military targets	Vals, 17.11.2026 (Jaff, Ozerov): 14 models in the drone task and 15 in the satellite task, three scenarios (two based on the war in Ukraine, one on Iran). No prompt was blocked by a classifier. The authors: "We did not establish how effective these models would be in a deployed weapons system." [8]	Confirmed. CORRECTION not all scenarios concern Ukraine.

3 Verification of key data

CONT. · ROWS 10–13

Claim	What the source says	Assessment
The Auterion contract	33,000 kits, about USD 50 million, a Pentagon contract, delivery announced by the end of 2025. The operator still selects the target. [13]	Confirmed as a contract and an announcement. Delivery: not verified. CORRECTION the contract was funded by the Pentagon, and the sources are the FT, The Register and DroneLife.
DELTA cyber-security audit	MoD, 09.03.2026: more than 160 requirements, no deviations. The Vezha module and the Avengers platform were covered. The assessment took about two months. The auditor was not disclosed as a specific company in the materials checked. [12]	Confirmed as an MoD statement.
ICRC principles	Whether a system is an autonomous weapon system (AWS) is determined by the function of selecting and applying force, not by the presence of AI. Autonomy in navigation or in engaging previously specified targets does not in itself make a system an AWS. A person's status can change within seconds. Recommendations include a ban on AWS against people and a requirement for effective oversight. [11]	Confirmed.
Microsoft: AI in influence operations	MTAC, April 2024: AI-enhanced content is more influential than fully generated content, and audio more than video. In the October 2024 report Storm-1679 used generative AI "with limited effect". Both findings concern operations targeting the US elections, not audiences of the war in Ukraine. [16]	Confirmed within scope. 2025 report: not verified.

3 Verification of key data

CONT. · ROWS 14–15

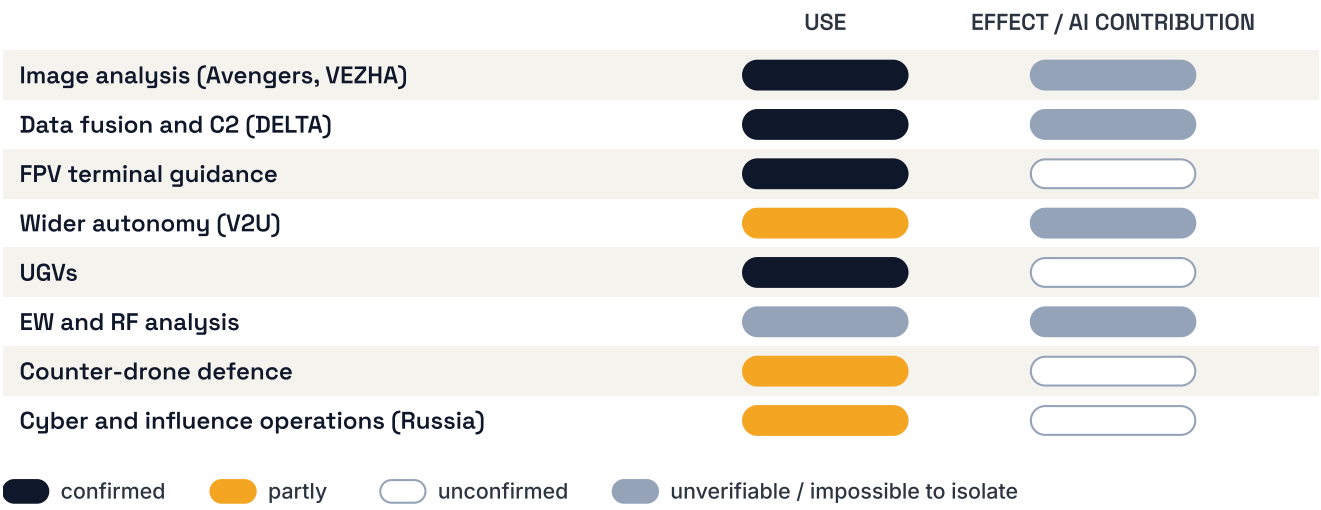
Claim	What the source says	Assessment
Exercise Combined Resolve 26-07	According to the brigade command's account (Defence Blog, via United24 Media), the Ukrainian team from the 412th "Nemesis" brigade caused about 70% of the US brigade's initial simulated losses over three days, and by the third day the share fell below 30%. The WSJ covered the exercise separately. [15]	Confirmed. An exercise result, not combat losses.
MWI: target hit on 43% of flights, 20–30% once missions not flown are included	The Modern War Institute article (Andreasen, 23.11.2026) could not be opened. The figures came from an earlier version of the report.	Not verified. Not to be treated as established fact.

Note on the last row

The figures from the Modern War Institute article are not used in this report as the basis for any conclusion. They remain in the table solely to make clear that they have not been confirmed.

4 Map of applications

STRENGTH OF EVIDENCE: USE VERSUS MEASURABLE EFFECT



Synthesis of the "Assessment" column in the table below and on the next page. No function has a publicly measured effect attributable to AI.

Function	What is documented	What is missing	Assessment
Image analysis (Avengers, VEZHA)	Detection of equipment in video from drones and fixed cameras, labelling standards, a separate validation set. [1][2]	False-alarm and miss rates by sensor, weather and object class.	Confirmed. Effect: impossible to calculate.
Data fusion and C2 (DELTA)	Integration of data from multiple sources, cloud and AI. Ifri describes DELTA and Kropyva as tools shortening the time from sensor to shot. [10]	A breakdown of what AI does versus ordinary integration.	Confirmed. AI contribution: impossible to isolate.
FPV terminal guidance	Maintaining the approach after link loss, target selected by a human. [5][6]	Share of missions completing the task, behaviour after link loss.	Confirmed. Universal increase in hits: unconfirmed.

4 Map of applications

CONT.

Function	What is documented	What is missing	Assessment
Wider autonomy (V2U)	Captured units, change in communications configuration, battlefield accounts. [3]	Logs, code, target-selection rules, error balance.	Partly confirmed.
UGVs	Ifri: testing and adaptation phase. Constrained by communications and the “last mile”. AI is used mainly for target tracking and fire correction, not for autonomous engagement. [10]	Share of missions completed without intervention.	Use confirmed. Autonomy: limited.
EW and RF analysis	EW is a continuous software contest at the tactical level. [10]	Data distinguishing ML from classical signal processing.	Unverifiable for claims about the AI contribution.
Counter-drone defence	Acoustic and thermal sensors, radar and data fusion, cueing for interception assets. [10]	An isolated AI contribution to defence effectiveness.	Partly confirmed.
Cyber and influence operations (Russia)	Anthropic’s September 2026 report describes Russia-linked actors using Claude, among other things for attacks on Ukrainian and European targets. In the espionage campaign humans retained direction, including decisions on targets. [17]	Independent verification beyond the model provider’s report.	Partly confirmed.

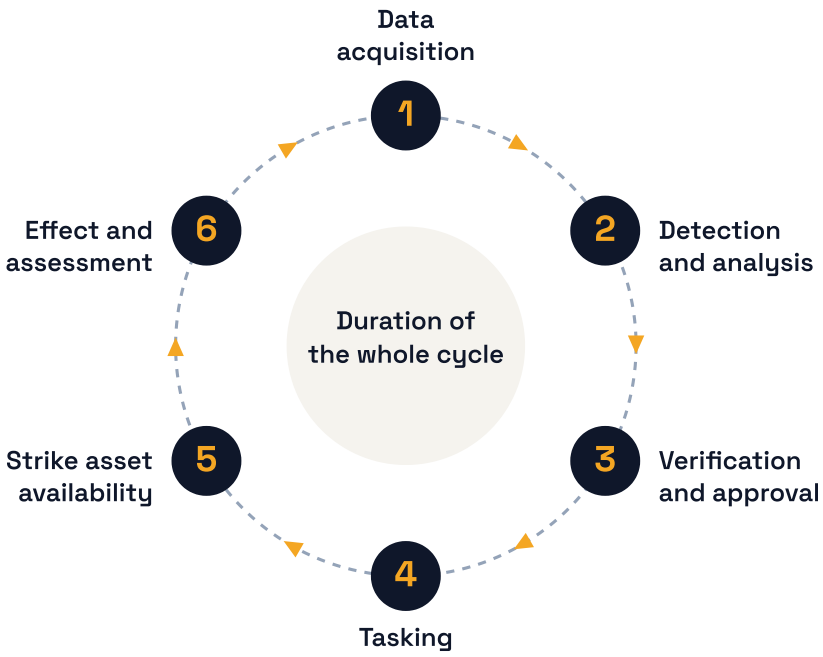
What we do not know about the models

Public material usually discloses the class of function, not the architecture, weights version, training set or validation procedure. CSIS describes the Ukrainian approach as training small, specialised models on small datasets and integrating them with various platforms, which is different from running a large language model on every drone. [2]

5 Analysis

5.1 The sensor-to-shooter loop and the denominator problem

AI can shorten the loop by organising data, and organising is the very mechanism that speeds it up. The difficulty lies in attributing the outcome. The duration of the whole cycle is made up of six stages: data acquisition, detection and analysis, verification and approval, tasking, availability of a strike asset, and achieving and assessing the effect.



Six stages of the sensor-to-shooter cycle. AI can speed up the organisation of data in several of them, but the outcome depends on the slowest link.

Ifri describes AI as a process accelerator rather than a decision-maker, and attributes the shortening of the cycle from days to minutes to the whole toolset, not to a single module. [10]

5.1 The denominator problem (cont.)

The same “effectiveness” can be calculated in four ways: against published videos, against activations of guidance, against all launches, or against all assigned tasks, including missions that could not be flown.

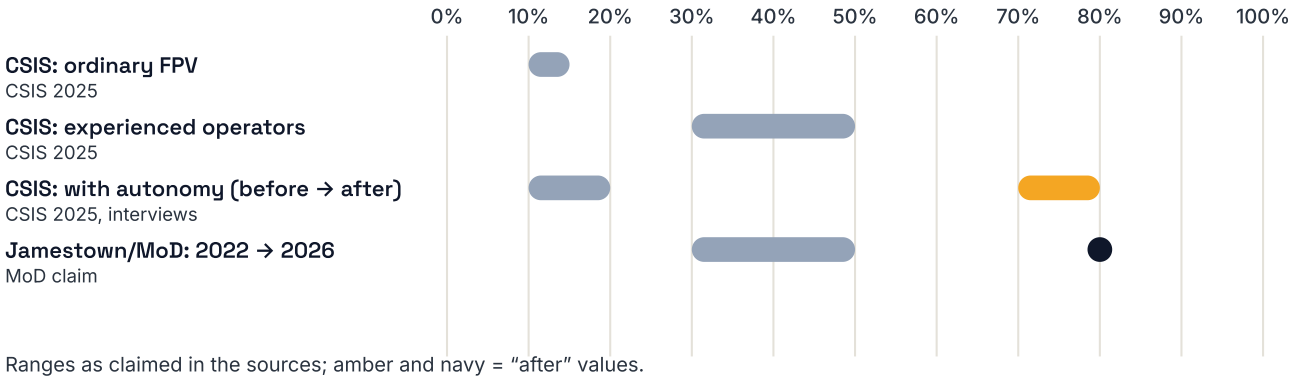
FOUR DENOMINATORS FOR THE SAME “EFFECTIVENESS”

Published videos	narrowest set: successful strikes get shown
Guidance activated	only flights in which the system took over the target
All launches	also flights ending in loss of the drone or link
All assigned tasks	also missions that could not be flown

The wider the denominator, the lower the “success” rate for the same events.

The figures from different sources look like this:

Source	Value	What it measures
CSIS, March 2025 [2]	10–15% ordinary FPV, 30–50% experienced operators, 10–20% to 70–80% with autonomy	Interviews, no defined denominator
Jamestown, 2026 (citing the MoD) [14]	30–50% at the start of the war, 80% in 2026	MoD claim, no methodology



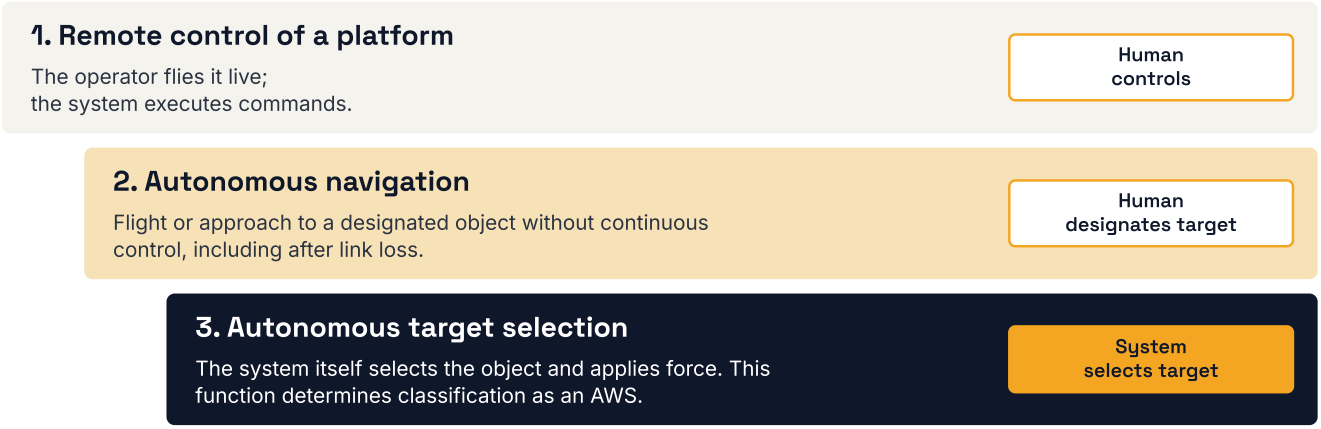
Ranges as claimed in the sources; amber and navy = “after” values.
This is not a single data series.

Conclusion

Values from 10% to 80% describe different events and different periods, and some are claims by the same party. They must not be combined into a single increase “thanks to AI”.

5.2 Autonomy: function by function

The debate on “autonomy” conflates three things: remote control of a platform, autonomous navigation and autonomous target selection.



Three different capabilities hidden under the single word “autonomy” (cf. the ICRC approach [11]).

The ICRC points out that classification is determined by the function of selecting and applying force, not by the presence of AI, and that autonomy in navigation or in engaging a previously designated target does not in itself make a system an autonomous weapon. [11] In June 2026 CSIS goes further, arguing that lethal decisions are settled not by what flies but by the software that controls it, and that drones carry out about 80% of strike missions on both sides. [4]

The human role is not binary. In every case one has to ask separately:

- who selects the specific object,
- who authorises the use of force,
- whether a human can monitor the rest of the flight,
- whether they can abort it,
- what happens after link loss.

A statement that an operator approved the target does not prove continuous oversight or the ability to call off the attack once the link is cut. This limitation follows directly from the described capability to fly without contact. [5]

5.2 Autonomy: the V2U case (cont.)

The assessment of V2U has changed over time.



Sources: CSIS [2][3]; data on captured units largely Ukrainian.

In March 2025 CSIS stated that autonomy in the sense of the US definition was not yet present on the battlefield in Ukraine. In April 2026 it assessed V2U as a probably fully autonomous system, noting that later units lacked LTE modems. The same report describes group coordination as inference. [2][3] The absence of a link does not settle whether the target was selected in advance, and accounts of group behaviour are no substitute for reconstructing how the code works.

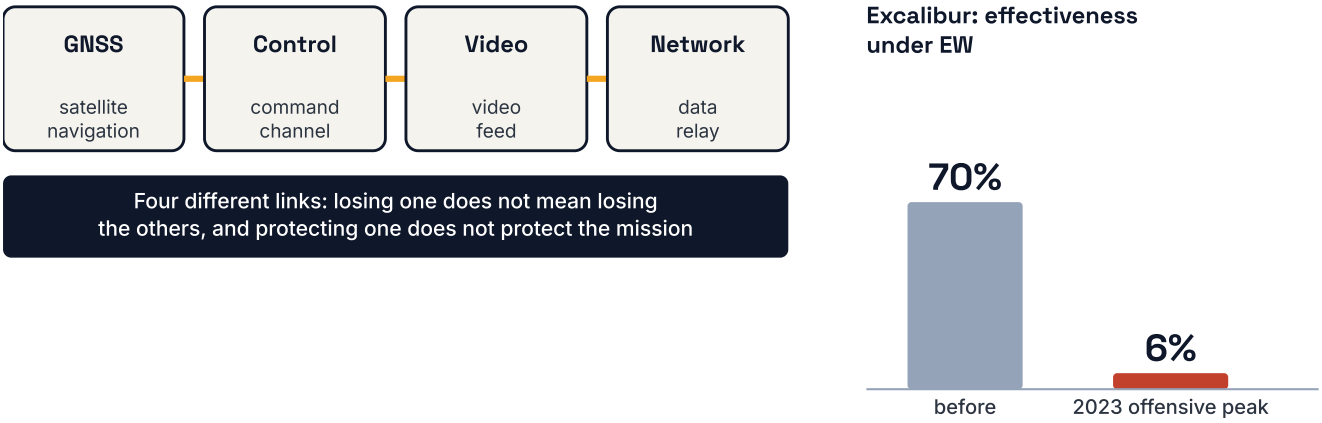
Thread: Anthropic report, September 2026

Anthropic's September 2026 report adds one more thread. According to Defense One, a Russian group used Claude to build a model small enough to run on a single-board computer and capable of selecting targets and issuing detonation commands with no human in the loop. The group had not yet deployed it in the field, but was testing it on real hardware. [17]

This is the model provider's claim, not an independent finding, and it concerns development, not combat use.

5.3 Electronic warfare as a constraint

GPS is only one of the GNSS systems, and losing satellite navigation is different from jamming of the control channel, loss of the video feed or being cut off from the network relaying the data. On-board image processing can make a specific phase of flight independent of external control. That does not prove the whole mission is resilient, because prior data, correct designation of the object and working sensors are still needed. [5][10]



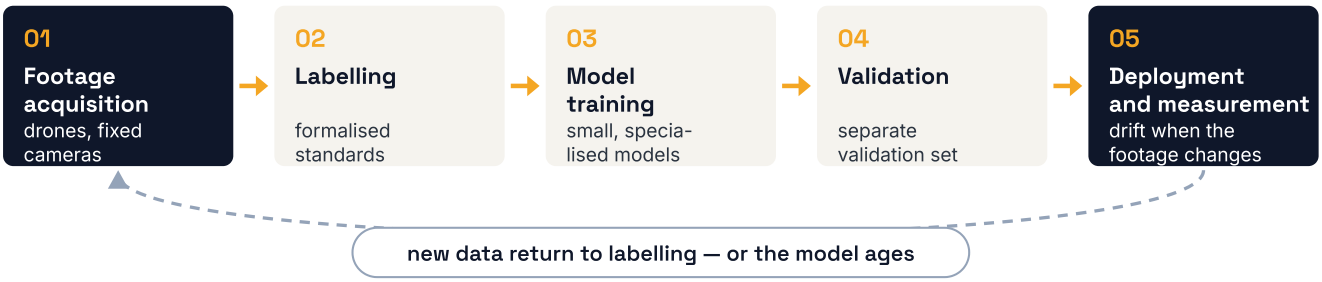
Left: links in the chain vulnerable to jamming. Right: fall in the effectiveness of Excalibur shells according to Ifri [10].

Ifri describes fibre-optic drones with 5–15 km of cable as a solution that removes the attack surface for jamming and shifts the fight from electronic suppression towards kinetic action. The authors also report that the effectiveness of Excalibur shells fell from 70% to 6% at the peak of the 2023 offensive owing to EW. Fibre optics are not AI, however, and do not make EW redundant, as the report shows its importance for other platforms, navigation and communications. [10]

There are no public, independent tests of the resilience of the systems in use to spoofing, visual deception and environmental change. There is, however, a documented example of model drift after a change of environment (described in Section 5.4). [2] The ICRC points to the vulnerability of systems to hacking, spoofing and countermeasures, but does not give the frequency with which these are exploited. [11]

5.4 Data, labelling and model drift

The most strongly documented technical conclusion is an unglamorous one.



The data pipeline as the real source of capability. Based on CSIS [2] and Ifri [10].

CSIS describes formalised labelling standards and a separate validation set. [2] According to CSIS, models trained on traffic cameras lost their usefulness when the nature of the footage changed, which shows that a model without current data ages along with the battlefield.

Ifri adds that tactical lessons about drones, AI and EW can become outdated within weeks. [10]

Conclusion

The quality of a capability is determined by the data pipeline and the pace at which it is renewed, not by the model itself.

5.5 Legal and model-safety risks

Visual recognition is not an assessment of legality. A model can detect a person, a uniform or a vehicle, but that does not settle whether an attack is permitted. The ICRC stresses that civilians are protected unless they directly participate in hostilities, and that the status of a person hors de combat, for example one who is surrendering, can change within seconds. According to the ICRC, autonomous weapons cannot reliably interpret such context. Legal obligations rest with the people who plan, approve and carry out an attack. [11]

What Vals AI showed. In the test, general-purpose models were given drone and satellite images and asked to identify targets. The drone scenes were synthetic, while the satellite images of the area around Minab in Iran and the drone footage from the war in Ukraine were real. The authors report that no prompt was blocked by a classifier, that refusals came from the models themselves, and that framing the task as a game often removed those refusals too. According to Vals, overall the school in Minab was identified in 70% of attempts, and Claude Opus 5.5 in 69% of standard attempts and 72% of attempts with the game framing. [8]

0
prompts blocked by a classifier
Vals AI [8]

70%
of all attempts identified the school in Minab
Vals AI [8]

69%
Claude Opus 5.5, standard attempts
Vals AI [8]

72%
Claude Opus 5.5, attempts framed as “a game”
Vals AI [8]

Scope of the study: 14 models in the drone task, 15 in the satellite task, three scenarios (two based on the war in Ukraine, one on Iran). [8]

5.5 Legal and model-safety risks (cont.)

What the study does not prove. The authors themselves recorded that they did not establish how effective such models would be in a deployed weapons system and did not assess targeting accuracy or compliance with the law of armed conflict. Nor were the models asked to make strike decisions, and commercial versions were tested. The result is a warning about safeguards and about the risk of mass-labelling footage for targeting systems, not a statistic of civilian casualties in Ukraine.

Two shortcuts are unjustified

- “the study proved that Ukrainian AI drones attack medics”,
- “a laboratory result is the percentage of mistaken attacks at the front”. [8][9]

A human in the loop is not an automatic guarantee. Real oversight requires information, time and the ability to intervene, and the ICRC recommends a requirement for effective oversight together with the ability to deactivate, as well as a ban on systems used against people. [11]

Effective oversight requires

- information,
- time,
- the ability to intervene and deactivate.

Audit ≠ accuracy

Meeting more than 160 cybersecurity requirements with no deviations is a different measure from classification accuracy. [12]

A cybersecurity audit is not an accuracy audit. The positive assessment of DELTA, VEZHA and Avengers covering more than 160 requirements does not prove error-free classification or resistance to manipulation. [12]

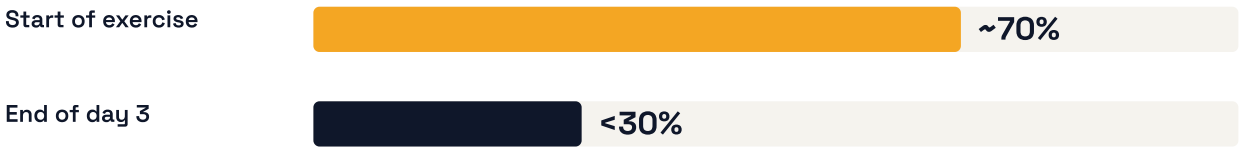
5.6 Influence operations: use is not effect

Microsoft documents the use of generative AI by Russian actors, but its findings are nuanced. In April 2024 MTAC wrote that AI-enhanced content is more influential than fully generated content, and audio more than video. In October 2024 it described Storm-1679's use of AI "with limited effect". [16] Both findings concern influence on the US elections and do not automatically carry over to audiences of the war in Ukraine or to 2026. What is confirmed is the use of the tools; a causal link between a specific deepfake and a political outcome, unit morale or the level of aid requires separate proof.

5.7 Exercises versus war

During exercise Combined Resolve 26-07 (April and May 2026, about 3,500 US soldiers) a Ukrainian team from the 412th "Nemesis" brigade caused, according to the brigade's account cited by Defence Blog, about 70% of the US armoured brigade's initial simulated losses in the first three days. By the end of the third day the share of losses caused by drones had fallen below 30%. The account estimates that about 60% of the successful defence resulted from discipline and procedures. [15]

Combined Resolve 26-07: drones' share of the US brigade's simulated losses



An exercise result, not combat losses. About 60% of the successful defence was attributed to discipline and procedures [15].

The US forces had standard electronic warfare and counter-drone assets, and their results improved during the exercise. [18] This demonstrates the importance of the whole team: persistent reconnaissance, specialised crews, direct communication and rapid adaptation. It is not an experiment comparing the same brigade with and without AI, and simulated losses are not combat losses.

6 Marketing versus evidence

Slogan	What can be defended
“Autonomous”	In public material this word is used for navigation, tracking a designated object, coordinating platforms and selecting targets. CSIS itself notes that Ukraine has no formal definition of autonomy or of an autonomous weapon, and that sources use “autonomous” and “unmanned” interchangeably. Without a description of the function, the label has no analytical value. [2]
“Three to four times more effective”	Improvement under specific conditions is credible. A constant multiplier independent of weather, crew and adversary has not been demonstrated. CSIS figures should be given as reported estimates. [2]
“10,000 times more effective”	A remark by a company founder quoted by NPR, with no defined metric. [6]
“Jamming-proof AI”	A specific on-board function may allow a specific phase of flight to continue after a specific link is lost. The claim that the drone cannot be jammed goes beyond the evidence. [5][10]
“Swarm”	Many drones used at once do not prove collective decision-making. Today’s technology more often corresponds to one operator controlling multiple platforms. [7]
“A cheap drone destroyed expensive equipment”	This may be a true description of a single event, but it is not an economic balance of a capability. Failed missions, reconnaissance, people, communications, servicing and maintenance have to be added. [10]
“A think-tank report proves effectiveness”	Not always. A report may correctly quote an interviewee’s unverified claim. One has to check whether the author measured it or collected accounts.

7 Panel claims tested against the evidence

The comparison covers claims made at the panel "AI on the Battlefield" ("SI na polu walki", Dell Technologies, Łódź, 24.11.2026) and shows which are supported by the sources checked. Attribution of the claims was established from the recording transcript, which does not separate speakers.

Claim	Support in the sources	Assessment
The problem is the training data, not the models themselves	CSIS: labelling standards, validation set, model drift after a sensor change. Ifri: AI as a process accelerator. [2][10]	Supported
A human approves the target; nobody has yet handed over autonomy of target selection	Consistent with Reuters and NPR for the Ukrainian cases described. Requires a caveat on Russia's V2U (CSIS: probably fully autonomous) and the Anthropic report. [3][5][6][17]	Supported with a caveat
Front-line operators do not understand how AI works	No study in the sources checked.	Own observation
For every hit there are 15 flights that do not reach the target	Implies about 6% effectiveness, below the ranges given by CSIS (10–15% ordinary FPV, 30–50% experienced operators). The sources give no such rate. [2]	Unconfirmed Label as own estimate
The Russians are poisoning data and models	No confirmation of models being poisoned with Russian data was found in the sources checked. What is documented is the use of commercial models by Russian actors (Anthropic report, September 2026). [17]	Unconfirmed in this form

7 Panel claims tested against the evidence

CONT.

Claim	Support in the sources	Assessment
Analysis of portable-toilet servicing data: three weeks of an analyst's work versus 15 minutes with AI	No independent measurement in the sources checked. The example illustrates a shorter analysis time, not its accuracy.	Anecdote, unsourced
We are reaching physics that AI cannot cheat	Ifri: EW, fibre optics, Excalibur effectiveness falling from 70% to 6%. Reuters: dependence on battlefield conditions. [5][10]	Supported
Poland cannot build a large model on its own; the European Union could	A strategic assessment, outside the scope of verification.	Opinion

Tally of the eight claims



Supported 2 · Supported with a caveat 1 · Unconfirmed 2 · Observation / anecdote / opinion 3

8 Recommendations for Poland and NATO

The points below are analytical recommendations drawn from the material, not a claim that there is a single proven model to copy.

1 **Data and integration first.** Consistent labelling and archiving of data, recording failures as well, common exchange formats, the ability to trace the origin of a designation, secure training and validation environments, and the ability to change supplier without losing data. Basis: the Avengers experience and the development of DELTA as an ecosystem. [2][10]

2 **Procure results under defined conditions, not an "AI" label.** The acceptance protocol should cover effectiveness against specified object classes, the number of false alarms, behaviour under uncertainty, operator workload, performance under degraded communications, results across seasons and maintenance cost. These are proposed criteria, justified by data drift and the rapid obsolescence of solutions. [2][10]

3 **Describe autonomy by function.** In requirements and reports, separate navigation, target tracking, target selection and authorisation of the use of force, in line with the ICRC approach, and require the ability to intervene and deactivate. [11]

4 **Do not build counter-drone defence on a single sensor type.** Ifri describes a detection ecosystem covering acoustics, thermal imaging, radar and data fusion, and notes that against fibre-optic and autonomous drones EW loses effectiveness and physical interception returns. The results of each detector therefore have to be assessed for specific threat classes. [10]

8 Recommendations for Poland and NATO CONT.

5 **Treat general-purpose models as a supply-chain and misuse risk.** The Vals results show that providers' safeguards do not guarantee refusal in tasks close to targeting, so procedures should assume verification of outputs and a limited role for them in the decision chain. [8][9]

6 **Do not copy statistics out of their context.** What transfers are principles: integration, rapid feedback, testing against an adaptive adversary, organisational resilience and sustaining capability. What does not transfer without further research are local hit rates, the economics of a section of the front and assumptions about the presence of civilians. [10]

7 **Scale in stages, without a single "readiness" date.** Pilots of analytical functions, validation with users and recording of errors, integration with procedures and training, scaling once a repeatable benefit has been shown, and continuous reassessment after changes to models and the environment. This is a recommended sequence, not a calendar forecast.



Recommended implementation sequence (point 7). A sequence, not a calendar.

9 Ten questions without a reliable public answer

1. What share of drone missions on both sides actually uses ML, rather than just an autopilot or classical tracking?
2. How does the effectiveness of the same team and platform change when AI is added under comparable conditions?
3. What are the full statistics of false and missed Avengers detections by sensor, weather and object class?
4. How many missions with automatic guidance end with the task accomplished, rather than mere contact with the object?
5. How often does the system behave differently from what the operator expected after link loss?
6. In what share of Russian V2U missions does the system select the target, and in what share was it specified in advance?
7. Is there an independently verified account of civilian harm resulting from autonomous target selection in this war?
8. What is the real cost of a successfully accomplished task once attempts, crews, data, training and maintenance are included?
9. How much of the claimed saving in operator work remains once oversight, quality control and retraining are included?
10. How quickly does the advantage of a given model version fade after the adversary changes tactics and countermeasures?

10 Scope of verification and limitations

Checked in the sources

Ukrainian MoD (Avengers, DELTA), CSIS (March 2025, April 2026, description of the June study), Reuters via The Star, NPR via KPBS, RUSI (swarms), Vals AI and TechCrunch, Ifri (in a summary of excerpts), the ICRC, accounts of the Auterion contract, accounts of Combined Resolve, Microsoft MTAC pages, and the RUSI, Euronews and Defense One accounts of the Anthropic report.

Not verified

The Modern War Institute article (the 43% and 20–30% figures), the Microsoft Digital Defense Report 2025, the Military Review article on integrating autonomous weapons, and RBC-Ukraine analyses. The full text of the June 2026 CSIS study could not be downloaded, so its claims are given according to the description on its web page and the Small Wars Journal review.

AI tools and provider caveat

Prepared with the help of AI tools. Two sources concern models from the provider of one of those tools: the Vals AI study and Anthropic's September 2026 report. They were treated like all the others, and results unfavourable to that provider are also given.

Origin of the figures

Some descriptions of the Vals study come from secondary accounts (TechCrunch, media write-ups). Where the original Vals text gave data, its wording was used. Figures from Ifri are given according to a summary of excerpts from the report.

Sources

[1] Ministry of Defence of Ukraine, statement on 12,000 targets detected per week, 23.09.2024.

[2] CSIS, “Ukraine's Future Vision and Current Capabilities for Waging AI-Enabled Autonomous Warfare”, March 2025. https://www.csis.org/analysis/ukraines-future-vision-and-current-capabilities-waging-ai-enabled-autonomous-warfare

[3] CSIS, analysis of the Russian V2U system, April 2026.

[4] K. Bondar, M. Mande, CSIS, “Defining Autonomy: Why Software, Not Drones, Will Decide the Next War”, June 2026 (according to the CSIS web page and the Small Wars Journal review of 16.06.2026). https://www.csis.org/analysis/defining-autonomy-why-software-not-drones-will-decide-next-war

[5] Reuters, “Ukrainian drone pilots look to AI for battlefield edge”, 29.11.2025 (The Star version).

[6] NPR, report on AI drones in the war in Ukraine, 29.10.2026 (KPBS version).

[7] R. Tollast, RUSI, commentary on drone swarms, 16.07.2026.

[8] G. Jaff, A. Ozerov, Vals AI, “Frontier models have few guardrails around autonomous weapons”, 17.11.2026. https://www.vals.ai/blogs/autonomous-warfare

[9] TechCrunch, report on the Vals AI study. https://techcrunch.com/2026/11/17/ai-models-can-be-tricked-into-identifying-civilians-as-military-targets-by-telling-them-its-a-game/

[10] Ifri, É. Tenenbaum et al., report on MilTech in the war in Ukraine, February 2026. https://www.ifri.org/sites/default/files/2026-02/ifri_tenenbaum_et_al_miltech_war_ukraine_2026.pdf

[11] ICRC, “Autonomous Weapon Systems and International Humanitarian Law: Selected Issues”, October 2025. https://www.icrc.org/sites/default/files/media_file/2025-10/ICRC-Position_Paper-Autonomous_Weapon_Systems_and_IHL-Selected_issues_Oct2025.pdf

Sources CONT.

[12] Ministry of Defence of Ukraine, statement on the independent cybersecurity assessment of DELTA, 09.03.2026.

[13] The Register, FT and DroneLife, reports on the delivery of 33,000 Auterion kits, July 2025.

[14] Jamestown Foundation, “Ukrainian Military Offers Lessons Learned to NATO, Part Three”.
https://jamestown.org/ukrainian-military-offers-lessons-learned-to-nato-part-three/

[15] United24 Media, report on Combined Resolve 26-07 (via Defence Blog, 20.10.2026).
https://united24media.com/world/ukraines-drone-veterans-hit-a-us-tank-brigade-with-a-70-casualty-lesson-in-modern-warfare-23646

[16] Microsoft MTAC, 17.04.2024 and 23.10.2024.
https://blogs.microsoft.com/on-the-issues/2024/04/17/russia-us-election-interference-deepfakes-ai/

[17] RUSI, “From Human Proxies to Autonomous Adaptation”; Euronews, 16.09.2026; Defense One, September 2026 (on the Anthropic report).
https://www.rusi.org/explore-our-research/publications/commentary/human-proxies-autonomous-adaptation-next-phase-russias-hybrid-activity

[18] SAN, report on Combined Resolve (via the Wall Street Journal).
https://san.com/cc/ukraines-drone-war-offers-hard-lesson-for-us-armory/
As of 25.11.2026: the address redirects to the site’s home page.

Source numbering in this edition is continuous (1–18); in the working draft, items 14–18 were numbered 15–19. Items without an address are given as in the source material.

MATTHEW HALABA.

AI on the Battlefield in Ukraine. What open data confirms, and what it cannot show. Analytical report, as of 25.11.2026.

Matthew Halaba · halaba.online · matthew@halaba.online