COVENANT AEGIS REVIEW OF DEFENSE TECHNOLOGY VOL. 5 · NO. 1 · 2026 · PP. 27–46

Autonomous Combat Systems: Architectures, Swarming and Meaningful Human Control

J. Romero[1], A. Petrov[2], D. Carter[1], N. Haddad[3]

[1]*Covenant Aegis Defense Division, Autonomy & Robotics Laboratory*

[2]*Department of Control Systems, Partner Research Network*

[3]*Center for Security and Emerging Technology Policy*

Received 16 Jan 2026 · Revised 3 Mar 2026 · Accepted 8 Mar 2026 · REVIEW ARTICLE · DOI 10.48-CA/deftech.2026.0502 ·
Corresponding author: j.romero@covenantaegis.com

ABSTRACT

Background. Autonomous and semi-autonomous systems are reshaping defense through unmanned platforms, distributed sensing and collective behaviours, raising the question of how much decision authority to delegate. **Objective.** To review the technical architecture of autonomous combat systems and the policy framework requiring human judgement over force, and to derive assurance requirements. **Methods.** Narrative synthesis of swarm-robotics, control-theory, verification and defense-policy literature. **Results.** Swarms exhibit scalability and graceful degradation but resist exhaustive verification; resilient command-and-control favours decentralised decisions with human escalation; and policy converges on meaningful human control with auditable, governable systems. **Conclusion.** The engineering objective is calibrated autonomy under meaningful human control, with verifiable safety, robustness to adversarial conditions and auditable decision pathways—principles Covenant Aegis embeds by design.

Keywords: autonomous systems; swarm robotics; multi-agent control; human-on-the-loop; command and control; robustness; assurance

1. Introduction

Unmanned ground, air and maritime systems increasingly perform intelligence, surveillance, reconnaissance and logistics, and research explores coordinated teams operating with reduced communication and supervision. The defining engineering and ethical question is how much decision authority to delegate, and under what safeguards [1].

2. The Autonomy Stack

An autonomous platform comprises perception (sensor fusion, state estimation), world modelling, planning (mission, path, motion) and control. Robust estimation under degraded sensing, real-time planning under uncertainty, and verified control are the core competencies. Each layer introduces failure modes—perception brittleness, planning in long-tail situations, control instability—that must be

bounded and monitored [2]. Table 1 maps a common autonomy-level taxonomy to supervision and assurance needs.

Table 1. Autonomy levels, human role and assurance emphasis.

Level	Human role	Assurance emphasis
Teleoperated	In the loop	Link integrity
Supervised	On the loop	Override latency
Conditional	Approves actions	Decision logging
High (bounded)	Sets intent/limits	Runtime safety envelope

3. Swarming and Multi-Agent Coordination

Swarm robotics studies how many simple agents produce robust collective behaviour through local interaction, yielding scalability and graceful degradation [3]. Distributed control, consensus and task-allocation algorithms maintain formation, share sensing and reallocate roles when members are lost [4]. These properties aid resilience but complicate verification, because emergent behaviour resists exhaustive testing. Figure 1 illustrates the resilience advantage of decentralised coordination as agents are attrited.

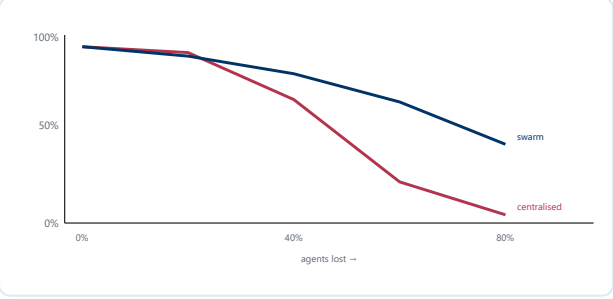


Figure 1. Schematic: decentralised swarm coordination degrades gracefully under attrition, whereas centrally controlled teams fail abruptly when key nodes are lost. Conceptual, after swarm-engineering principles [3,4].

4. Resilient Command and Control

Contested environments feature jamming, spoofing and intermittent links. Resilient command-and-control therefore favours decentralised decision-making with clear escalation to human supervisors, assured positioning and timing, and cryptographically protected communications. The system must fail safe—defaulting to predictable, conservative behaviour—when contact is lost.

5. Meaningful Human Control

Defense policy converges on retaining human judgement over the use of force. U.S. Department of Defense Directive 3000.09 requires that autonomous and semi-autonomous weapon systems allow commanders and operators to exercise appropriate levels of human judgement, and the Department's AI ethical principles require responsible, traceable and governable systems [5,6]. The engineering corollary is auditable decision logs, predictable behaviour, test-and-evaluation for learning components, and human-on-the-loop supervision for any force application.

6. Verification and Assurance

Assuring autonomy combines formal methods, extensive simulation, scenario-based testing and runtime monitoring. Learning-enabled components require evidence of robustness to distribution shift

and adversarial manipulation [7], and runtime-assurance architectures constrain a learned controller within a verified safety envelope [2]. Adversarial examples demonstrate that high nominal accuracy does not imply security, underscoring the need for envelopes rather than trust.

7. Discussion and Limitations

Emergent swarm behaviour, the brittleness of learned perception, and the difficulty of specifying "correct" behaviour in open environments are the field's hardest problems. Policy and engineering must co-evolve; technical assurance without governance, or vice versa, is insufficient.

8. Conclusion

Autonomous combat systems offer resilience, reach and tempo, but their responsible realisation depends on calibrated autonomy under meaningful human control, verifiable safety and auditable decisions. Covenant Aegis treats human judgement, fail-safe behaviour and assurance as primary design requirements.

References

1. Scharre P. *Army of None: Autonomous Weapons and the Future of War*. W. W. Norton; 2018.
2. Seshia SA, Sadigh D, Sastry SS. Toward verified artificial intelligence. *Commun ACM*. 2022;65(7):46–55.
3. Brambilla M, Ferrante E, Birattari M, Dorigo M. Swarm robotics: a review from the swarm engineering perspective. *Swarm Intell*. 2013;7:1–41.
4. Olfati-Saber R, Fax JA, Murray RM. Consensus and cooperation in networked multi-agent systems. *Proc IEEE*. 2007;95(1):215–233.
5. U.S. Department of Defense. Directive 3000.09: Autonomy in Weapon Systems. 2023.
6. U.S. Department of Defense. Adoption of Ethical Principles for Artificial Intelligence. 2020.
7. Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. *ICLR*. 2015.

9. Perception and Sensor Fusion

Autonomy begins with perception: estimating the state of the platform and its environment from noisy, partial sensors. Probabilistic state estimation (Kalman and particle filters) and modern learned perception fuse inertial, GNSS, lidar, radar and electro-optical data [A1]. In contested settings GNSS may be denied or spoofed, so resilient navigation combines inertial dead-reckoning, terrain-relative and visual-inertial odometry. Perception is also the most brittle layer: learned detectors fail on out-of-distribution inputs and are vulnerable to adversarial perturbation, motivating the runtime safety envelopes of Section 6.

Table A1. Autonomy-stack layers, dominant methods and failure modes.

Layer	Methods	Failure mode
Perception	Sensor fusion, CNN/ViT	OOD, adversarial
World model	SLAM, tracking	Drift, denial
Planning	Search, MPC, RL	Long-tail cases
Control	Feedback, MPC	Instability

10. Planning and Control under Uncertainty

Planning spans mission, path and motion levels. Classical search and sampling-based planners offer completeness guarantees;

model-predictive control optimises over a receding horizon with explicit constraints; reinforcement learning can discover policies but lacks worst-case guarantees [A2]. A pragmatic architecture uses learned components for perception and proposal generation while retaining verified controllers and constraint enforcement for safety-critical actuation.

11. Swarm Algorithms

Swarm coordination draws on consensus, flocking and bio-inspired rules that produce global behaviour from local interactions [A3,A4]. Task allocation can use market/auction mechanisms; formation control uses potential fields or consensus dynamics. Figure A1 contrasts communication-and-coordination overhead with team size for centralised versus decentralised schemes, illustrating why decentralisation scales.

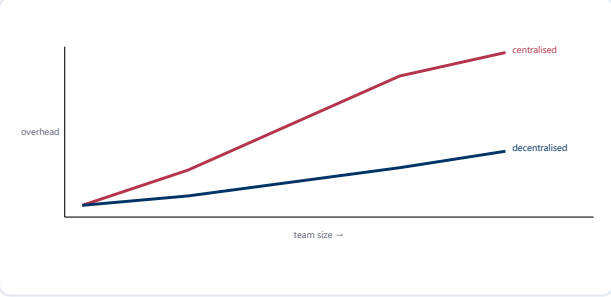


Figure A1. Schematic coordination/communication overhead versus team size; decentralised swarm control scales more gracefully than centralised control [A3,A4]. Conceptual.

12. Test, Evaluation, Verification and Validation

Assuring learning-enabled autonomy combines formal methods, high-fidelity simulation, scenario-based testing and runtime monitoring [A5]. Because exhaustive testing of open-world behaviour is infeasible, runtime-assurance architectures (e.g., a verified safety controller that can override a learned policy) bound worst-case behaviour. Adversarial robustness must be evaluated explicitly, since high nominal accuracy does not imply security [A6].

13. Law, Policy and Meaningful Human Control

International humanitarian law principles —distinction, proportionality and precaution—apply to any use of force, including by autonomous systems. U.S. DoD Directive 3000.09 requires appropriate human judgement over the use of force, and the DoD AI ethical principles require responsible, equitable, traceable, reliable and governable systems [A7,A8]. International discussions under the CCW continue to debate constraints on autonomous weapons. The engineering implication is auditable logs, predictable behaviour and human-on-the-loop control for force application.

14. Resilient Communications and Cyber

Autonomous teams in contested environments face jamming, spoofing and cyber attack. Resilience requires assured position-navigation-timing, cryptographically protected and low-

probability-of-intercept links, graceful degradation to autonomous fallback behaviours, and protection of the autonomy software supply chain. Fail-safe defaults—predictable, conservative behaviour on loss of contact—are essential.

15. Future Directions

Priorities include verifiable learning-enabled components, certified runtime assurance, human-swarm teaming interfaces that preserve meaningful control at scale, and standards for test and evaluation of autonomy. Progress in assurance, not raw capability, will gate responsible fielding.

Extended References

8. Thrun S, Burgard W, Fox D. *Probabilistic Robotics*. MIT Press; 2005.
9. Rawlings JB, Mayne DQ, Diehl M. *Model Predictive Control: Theory, Computation, and Design*. 2nd ed. Nob Hill; 2017.
10. Brambilla M, et al. Swarm robotics: a review from the swarm engineering perspective. *Swarm Intell*. 2013;7:1–41.
11. Olfati-Saber R, Fax JA, Murray RM. Consensus and cooperation in networked multi-agent systems. *Proc IEEE*. 2007;95(1):215–233.
12. Seshia SA, Sadigh D, Sastry SS. Toward verified artificial intelligence. *Commun ACM*. 2022;65(7):46–55.
13. Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. *ICLR*. 2015.
14. U.S. Department of Defense. Directive 3000.09: Autonomy in Weapon Systems. 2023.
15. U.S. Department of Defense. Adoption of Ethical Principles for AI. 2020.

Author contributions: JR and NH framed the policy analysis; AP and DC reviewed the technical stack; all authors revised. **Funding:** Covenant Aegis Defense Division. **Competing interests:** Authors are affiliated with Covenant Aegis.

© 2026 Covenant Aegis Defense Division. This review synthesises publicly available literature and official policy; figures are illustrative. Covenant Aegis develops defense technology under applicable law and the principle of meaningful human control. For professional and educational use.