En Jessica Yan, *Grammars of Asymmetric Visibility*. Montevideo (Uruguay): MAAT.

From Occupation to Optimization: AI, Military Language, and the Administrative Translation of Force.

Agustin V. Startari.

Cita:

Agustin V. Startari (2026). *From Occupation to Optimization: AI, Military Language, and the Administrative Translation of Force. En Jessica Yan Grammars of Asymmetric Visibility. Montevideo (Uruguay): MAAT.*

Dirección estable: https://www.aacademica.org/agustin.v.startari/243

ARK: https://n2t.net/ark:/13683/p0c2/fw9

From Occupation to Optimization: AI, Military Language, and the Administrative Translation of Force

How Predictive Systems Convert Violence into Operational Syntax

Author: Agustin V. Startari

Author Identifiers

- ResearcherID: K-5792-2016
- ORCID: https://orcid.org/0009-0001-4714-6539
- SSRN Author Page: https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7639915

Institutional Affiliations

- Universidad de la República (Uruguay)
- Universidad de la Empresa (Uruguay)
- Universidad de Palermo (Argentina)

Contact

- Email: astart@palermo.edu
- Alternate: agustin.startari@gmail.com

Date: August 20, 2026

DOI

- Primary archive: **10.5281/zenodo.22033285**
- Secondary archive: 10.6084/m9.figshare.33299976
- SSRN: Pending assignment (3rd Quarter 2026)

Language: English

Series: *Grammars of Asymmetric Visibility: AI, Imperial Power, and the Syntax of Responsibility*

Word count: 10,194 (main text)

Keywords: military AI; administrative translation of force; operational syntax; artificial intelligence; AI-enabled decision support systems; predictive targeting; occupation; Palestine; Gaza; Iran; surveillance; collateral damage; proportionality; risk management; security discourse; asymmetric visibility; responsibility loss; force optimization; civilian harm; computational discourse analysis.

Abstract

This article introduces administrative translation of force as a formal effect of AI-mediated military and security discourse. The concept describes a transformation in which violence, occupation, surveillance, targeting, or coercive control remains semantically present but is re-expressed through the operational vocabulary of risk management, target validation, collateral estimation, proportionality assessment, escalation control, security response, or optimization. The claim is not that these categories are inherently illegitimate. Modern armed forces require procedures for target development, legal review, civilian-harm mitigation, and decision support, and international governance frameworks increasingly insist that human responsibility and judgment remain traceable when artificial intelligence is used in military decision-making. The analytical problem begins when procedural language changes the grammatical architecture through which force is recognized: actors become systems, decisions become thresholds, attacks become operations, civilian harm becomes an estimate, and responsibility becomes distributed across models, scores, intelligence pipelines, and authorization chains.

The article uses Gaza-related reporting on the reported military use of AI-enabled decision-support systems as its principal empirical anchor, while Iran-Israel and United States security discourse function as comparative environments for examining the language of precision, pre-emption, targets, threat, escalation, and operational necessity. It does not infer undisclosed technical use from political language, and it treats contested claims about specific systems as reported claims rather than established technical facts. Building on prior work on responsibility loss, asymmetric visibility, and political-subjecthood reduction, the paper argues that operational syntax can preserve factual description while weakening the linguistic traceability of force, decision, civilian consequence, and institutional agency (Startari, 2025a, 2026a, 2026b, 2026e).

Methodologically, the paper proposes Operationalization Density (OD) and a Force Optimization Index (FOI). OD measures how frequently clauses about force are structured through administrative, predictive, proportionality-based, or optimization-oriented categories. FOI measures the combined degree of action nominalization, system-

mediated responsibility, civilian-harm compression, operational-necessity framing, and decision-chain opacity. The framework is explicitly falsifiable: if operationalization occurs at comparable rates across actors, source genres, and power positions, or if source texts already contain the same structure, the stronger geopolitical interpretation must be narrowed. AI ethics and military-AI governance should therefore ask not only whether systems are accurate, explainable, or lawful, but whether the language surrounding their outputs preserves the chain through which force is chosen, applied, experienced, and attributed.

1. Introduction: When Violence Becomes Procedure

Political violence does not have to disappear from language in order to become less politically legible. It can remain visible as a sequence of targets, operations, effects, estimates, warnings, responses, risk scores, threat assessments, or proportionality calculations. In that form, the event may be described with technical precision while the grammar of force becomes progressively administrative. A strike is represented as target prosecution. A person becomes a confidence-ranked candidate. A home becomes a targetable location. A civilian death becomes an estimated incidental effect. A command decision becomes the satisfaction of a threshold. A coercive act becomes a security response. The vocabulary may be accurate within an institutional workflow and still alter what kind of event the sentence allows a reader to perceive.

This paper calls that transformation the **administrative translation of force**. The term refers to the AI-mediated process through which military violence, occupation, surveillance, targeting, or coercive control is redescribed primarily as operational procedure, prediction, validation, optimization, proportionality management, or risk administration. Translation is used deliberately. The claim is not that the underlying act is necessarily concealed. Rather, the act is re-encoded into a different grammatical ontology. Instead of actor-action-object, the dominant structure becomes system-variable-output. Instead of a commander choosing among uncertain courses of action, a recommendation is generated, a target is validated, a risk is scored, a collateral estimate is produced, and an operation is authorized. The underlying human decision may remain legally indispensable, yet it becomes linguistically less central.

This question has become more urgent as artificial intelligence is incorporated into military decision-support systems. The International Committee of the Red Cross (ICRC) describes AI-enabled decision-support systems as computerized tools that aggregate data and provide analyses, recommendations, or predictions to human decision-makers, including in decisions concerning the use of force (International Committee of the Red Cross [ICRC], 2025). The same body stresses that international humanitarian law (IHL) requires human commanders and combatants to make legal determinations and remain accountable for them. Military and alliance policy documents similarly emphasize responsibility, traceability, reliability, and human judgment. The United States Department of Defense (DoD), for example, frames AI adoption around "decision advantage," speed, battlespace awareness, and resilient kill chains while also maintaining responsible-AI principles that require human responsibility and traceability (U.S. Department of Defense, 2023a, 2023b). NATO's revised strategy likewise combines accelerated adoption with responsibility, accountability, explainability, traceability, reliability, governability, and bias mitigation (North Atlantic Treaty Organization [NATO], 2024).

These texts establish an important tension. Military AI is explicitly designed to transform information into actionable recommendations more rapidly and at greater scale, but the legal and ethical frameworks governing it require that human judgment not be displaced by the system. The tension is usually analyzed as a problem of automation bias, reliability, explainability, proportionality, precautions, or meaningful human control (Blanchard & Bruun, 2025; Dorsey, 2025; Dorsey & Bo, 2026; Woodcock, 2023). This article isolates an adjacent problem: what happens in language after the decision environment has been computationally reorganized? If the system's output is expressed as a target, score, recommendation, or validated option, and if downstream reports preserve those categories while omitting the human decisions that transformed the output into force, then responsibility can be weakened without any formal transfer of legal responsibility to the machine.

The problem extends a sequence developed in earlier work. The passive voice in AI-generated language can weaken visible agency by making the actor optional (Startari,

2025a). The appearance of objectivity can be produced through impersonal form rather than evidentiary neutrality (Startari, 2025b). Syntactic authority can be generated by structural coherence even where a responsible speaker is absent (Startari, 2025d). In conflict discourse, the humanitarian passive describes suffering while weakening the grammatical path to responsible actors (Startari, 2026a). Asymmetric visibility describes cases in which multiple actors are mentioned while grammatical agency is distributed unequally (Startari, 2026b). Sanctions discourse can translate externally imposed coercion into pressure, restriction, or economic condition while weakening the visibility of the sanctioning actor (Startari, 2026c). Platform governance can perform a related conversion when restrictive effects are described as policy enforcement while the enforcing institution recedes from the grammar of suppression (Startari, 2026d). The immediately preceding paper in this series extends the problem to political subjecthood, asking whether subordinated societies remain visible primarily as risk, humanitarian need, instability, or objects of administration (Startari, 2026e). The present paper moves one step downstream: once people, territories, and organizations have become administratively legible objects, how does military AI language convert coercive action against them into operational procedure?

The research question is therefore: **How does AI-mediated military and security language transform force, occupation, surveillance, and coercive control into management, optimization, prediction, or technical procedure, and under what conditions does that transformation weaken the traceability of human decision and responsibility?**

The central thesis is conditional. AI-mediated military discourse may transform force into administrative syntax by preferentially describing violent or coercive action through operational, predictive, proportionality-based, security-oriented, and optimization-centered categories. The result is not necessarily denial. Indeed, the system may preserve substantial factual information: a target existed, an assessment occurred, an effect was estimated, a strike was authorized, and civilian harm followed. The change lies in the grammatical route through which those facts are connected. Force becomes a managed

output of a process rather than an action chosen by identifiable institutions and individuals.

This formulation requires several boundaries. First, **operational language is not itself evidence of illegitimacy**. Militaries require technical vocabulary; legal advisers require terms of art; planners require abstraction; and civilian-harm mitigation depends on structured procedures. Second, **proportionality assessment is a legal obligation, not a euphemism by definition**. The analytical question is whether the existence of an assessment is represented as if it were itself proof of legality, or whether the judgment, uncertainty, assumptions, and consequences remain visible. Third, **not every security claim is fabricated and not every target is misclassified**. The framework measures representational structure, not the truth of every underlying intelligence claim. Fourth, **specific allegations about AI systems must be sourced and qualified**. Where public reports describe systems such as the Gospel, Lavender, or Where's Daddy, this paper treats their characteristics as reported or alleged unless corroborated by authoritative documentation.

Gaza provides the strongest documented case for examining this translation because United Nations reporting has directly addressed the reported use of AI-enabled systems in target generation and has raised concerns about the speed, scale, and adequacy of human review (United Nations Human Rights Council, 2024, 2026). Iran provides a different comparison. The June 2025 Israel-Iran conflict generated a dense public vocabulary of precise strikes, pre-emption, imminent threat, nuclear facilities, targets, retaliation, and escalation, while IAEA reporting documented attacks on safeguarded nuclear facilities and the interruption of verification activities (International Atomic Energy Agency [IAEA], 2025; United Nations Department of Political and Peacebuilding Affairs [DPPA], 2025). The present paper does not infer military-AI use in those attacks. Instead, Iran functions as a comparative discourse environment for testing whether direct coercive relations are compressed into the administrative language of threat reduction and target management.

The methodological objective is not to declare that a phrase such as *target validation* or *collateral estimate* is politically suspect. It is to measure what such phrases replace, how

frequently they occur, which actors remain visible around them, and whether machine-mediated transformation increases their dominance relative to the source. The paper therefore proposes a clause-level coding system, Operationalization Density (OD), and a Force Optimization Index (FOI). These measures are designed to distinguish necessary military abstraction from systematic responsibility displacement. The paper does not manufacture corpus results. It specifies the conditions under which the theory can be confirmed, narrowed, or rejected.

2. From Digital Dehumanization to Operational Syntax

The preceding paper in this series distinguished **subject visibility** from **object visibility**. A society can remain intensely visible while appearing mainly as a humanitarian caseload, security risk, migration pressure, sanctions target, or field of instability rather than as a political subject capable of claim, decision, memory, resistance, sovereignty, or self-description (Startari, 2026e). The present paper begins where that reduction becomes operationally consequential. Once a population or territory is legible as a risk object, it becomes easier for an institutional system to process it through categories that belong to management rather than politics.

The transition can be represented as a sequence:

political subject -> risk category -> target candidate -> validated target -> estimated effect -> operational action

This sequence does not describe every military process and should not be read as a claim that political subjecthood is a prerequisite for lawful targeting. Its function is analytical. At each step, the represented entity acquires a narrower institutional role. A person who is initially described as a member of a political community may later appear as a node, association, profile, pattern, suspected operative, target candidate, or target. A neighborhood may appear first as lived territory, then as an operational environment, target-rich area, evacuation zone, or battlespace. A civilian population may appear first as rights-bearing persons and later as a density estimate, collateral presence, protected-object layer, or constraint in a strike calculation. Each transformation can serve a legitimate operational function. The question is whether the cumulative transformation

also narrows the grammar through which human and institutional responsibility remains recoverable.

This is the distinction between **object conversion** and **operationalization**. Object conversion concerns what a person, population, or territory becomes in discourse: risk, burden, threat, caseload, proxy environment, or administrative object. Operationalization concerns what force becomes once directed toward that object: validation, neutralization, risk reduction, target prosecution, collateral estimation, escalation management, or optimization. The first alters the represented status of the recipient. The second alters the represented status of the action.

Consider four constructed clauses:

The commander authorized an airstrike against the building after receiving intelligence that it was being used for military purposes.

The building was approved as a military target after intelligence review.

The target passed validation and collateral-estimation thresholds.

The system generated an actionable target within acceptable operational parameters.

All four can describe parts of the same institutional chain, but they allocate agency differently. The first clause names a human decision-maker and a coercive act. The second shifts the grammatical subject to the building and the category of target. The third makes the target and the threshold the principal actors of the sentence. The fourth foregrounds the system and treats the target as an output satisfying parameters. The progression is not from truth to falsehood. It is from a relational grammar of action toward a procedural grammar of eligibility.

The change matters because military decision support is increasingly built around precisely these intermediate objects. AI-DSS do not merely "know" a battlefield. They classify observations, rank patterns, recommend courses of action, estimate effects, identify anomalies, and sometimes generate proposed targets (ICRC, 2024, 2025). Dorsey and Bo (2026) emphasize that the important legal question is not confined to fully autonomous weapon systems; decision-support tools can influence core targeting functions even when a human formally remains in the loop. Blanchard and Bruun (2025) similarly distinguish AI-DSS from autonomous weapons while showing that both can

reshape the human role in targeting. The linguistic counterpart is that a system may leave legal authority formally human while making the machine-generated category the most visible subject in the operational record.

This paper therefore uses the term **operational syntax** for clauses in which force is represented primarily through institutional process variables rather than direct coercive relations. Operational syntax typically privileges nouns over verbs, criteria over actors, outputs over judgments, and system states over decisions. Its recurring forms include:

- nominalization: *targeting, validation, neutralization, escalation, degradation, mitigation*;
- threshold language: *acceptable risk, confidence level, collateral threshold, priority score*;
- procedural passives: *was validated, was cleared, was nominated, was assessed*;
- system subjects: *the model identified, the system recommended, the platform flagged*;
- necessity frames: *required, necessary, unavoidable, mission-essential*;
- optimization frames: *efficient, precise, minimized, accelerated, prioritized*;
- security-response frames: *response, containment, deterrence, threat reduction*.

None of these grammatical forms is inherently deceptive. Nominalization is indispensable to technical writing; passives can appropriately foreground the patient; scoring is necessary in many decision environments; and security planning cannot be written solely in ordinary verbs. The analytical issue is distribution and substitution. **What disappears when operational syntax increases?** If direct action verbs, human decision-makers, uncertainty markers, affected persons, and institutional authorizers remain present, operationalization may improve precision without reducing responsibility. If they disappear while procedural nouns and system outputs multiply, the same precision can create responsibility opacity.

This is where the concept differs from euphemism. Euphemism softens or replaces a socially charged expression. Administrative translation can occur even when no word is softened. *Target validation* may be the exact procedural name of a required step. *Collateral damage estimation* may be the correct technical term for an ex ante process. *Battle damage assessment* may be indispensable to operational review. The translation

effect arises not because the terms are inaccurate but because they alter the grammar of the event. The system no longer asks primarily who did what to whom; it asks whether a candidate object met criteria, whether expected effects fell below a threshold, and whether the operation advanced a mission objective.

A second distinction concerns temporality. Direct-force grammar tends to organize events around acts: *struck, detained, surveilled, destroyed, displaced, blocked*. Operational grammar tends to organize them around states and workflows: *under review, in the target set, within parameters, under monitoring, pending validation, subject to restriction*. This temporal shift makes coercion appear as an ongoing condition of administration rather than as a sequence of decisions. The effect is especially visible in occupation and surveillance, where control is often exercised continuously rather than through a single dramatic event. If movement restrictions, biometric screening, predictive watchlists, permit systems, or aerial surveillance are narrated only as security management, the political relation of control may be reduced to system maintenance.

Prior work on non-referential authority helps explain why this matters. Syntactic sovereignty describes how formal structure can generate the appearance of authoritative necessity even where agency and source are weak (Startari, 2025c). Null subjects of power identify institutional statements in which action survives while the actor recedes (Startari, 2025d). Compiled norms describe a related movement from interpretation toward executable form (Startari, 2025e). In military AI, these patterns converge: a target can become executable because it satisfies a sequence of formal conditions, and the language describing the sequence can make the conditions appear to act on their own.

The stronger political claim, however, requires comparison. If operational syntax occurs equally in descriptions of every military actor, every conflict, and every target population, then it may reflect a general property of military bureaucracy or AI summarization rather than asymmetric translation of force. If a source-controlled corpus shows that machine-generated text increases operationalization more strongly when describing dominant-power violence than when describing the violence of subordinated actors, the geopolitical interpretation becomes stronger. Conversely, if the source already

contains the same degree of operational abstraction, the transformation effect must be reduced.

The theoretical bridge can therefore be stated as follows: **digital dehumanization makes populations administratively processable; operational syntax makes force against administratively processed populations procedurally legible.** The first concerns who can appear as a subject. The second concerns what can appear as an action. Their combination creates a representational environment in which people are visible as objects and violence is visible as procedure.

3. The Administrative Translation of Force

Administrative translation of force operates through a set of recurrent transformations. These transformations are not unique to artificial intelligence; military and bureaucratic language long predates machine learning. What changes under AI mediation is speed, scale, recombination, and the possibility that intermediate computational outputs become stable linguistic objects across intelligence, targeting, reporting, summarization, legal review, and public explanation.

The first mechanism is **action-to-process conversion**. A finite action with an identifiable agent becomes an abstract process. *Forces struck the compound* becomes *target engagement occurred*. *Authorities restricted movement* becomes *movement-control measures were implemented*. *The military surveilled the neighborhood* becomes *persistent ISR coverage was maintained*. Nominalization does not eliminate the event, but it removes the grammatical requirement to state an acting subject. This mechanism extends the responsibility-loss effects previously identified in AI-generated conflict discourse (Startari, 2025a, 2026a). In a military-AI context, the process noun can also connect directly to a software workflow, making the institutional sequence appear more salient than the force itself.

The second mechanism is **actor-to-system substitution**. Human decision-makers are displaced by technical subjects: *the system identified*, *the model ranked*, *the intelligence pipeline assessed*, *the platform generated*, *the algorithm recommended*. This is not necessarily false. A system may indeed perform classification or recommendation. The displacement occurs when the sentence ends there. A model can recommend; it cannot

bear the same legal and institutional responsibility as a commander who authorizes force. The ICRC's 2025 submission is explicit that legal determinations remain human obligations and that AI-DSS must support rather than replace human decision-making (ICRC, 2025). A linguistically complete account therefore has two chains: **computational contribution** and **human authorization**. If only the first survives, procedural agency replaces accountable agency.

The third mechanism is **harm-to-variable conversion**. Civilian presence, injury, infrastructure destruction, displacement, or reverberating effects become quantities inside a decision space: *estimated civilian casualties, collateral effects, protected-site proximity, population density, expected damage*. Quantification can improve precaution and may be indispensable to civilian protection. The problem arises when the variable becomes the final representation of the people it measures. Dorsey (2025) describes a broader history of "quantification logics" in targeting and warns that computational models may narrow the contextual judgment required by proportionality. The representational counterpart is that civilian harm can become visible as a number while the lived and causal structure of harm becomes less visible as an event.

The fourth mechanism is **decision-to-threshold conversion**. A choice made under uncertainty becomes the passing or failing of a formal criterion. *The commander judged that the anticipated military advantage outweighed expected incidental harm* becomes *the strike remained within the collateral threshold*. The second clause may summarize a procedural conclusion, but it changes the locus of action. The threshold appears to authorize. The decision-maker appears to comply. This grammar is especially powerful because it converts contestable judgment into the appearance of mechanical necessity. In contexts where data are incomplete, probabilities uncertain, and civilian behavior dynamic, the apparent precision of a threshold can exceed the epistemic precision of its inputs (Dorsey, 2025; ICRC, 2024).

The fifth mechanism is **force-to-response framing**. Initiating, escalating, or sustained coercive action is represented through a relational category such as *response, deterrence, containment, protection, stabilization, pre-emption,* or *threat reduction*. These categories may be factually and legally relevant. A state can act in self-defense; an armed group can

initiate violence; deterrence can be a genuine strategic objective. The analytical issue is whether the response frame preserves the action that constitutes the response and the prior actions that constitute the asserted threat. If one side's force is repeatedly lexicalized as *response* and the other's as *attack*, the discourse can redistribute initiative before any explicit moral judgment appears.

The sixth mechanism is **sequence-to-dashboard compression**. Complex political and legal relations are reduced to a small set of operational indicators: target count, confidence score, threat priority, estimated collateral effect, mission completion, damage assessment, risk level, or escalation status. Dashboards are designed precisely to reduce complexity. In many contexts that is their virtue. Yet the same compression can narrow the range of facts available for deliberation. A dashboard can show that a target is “high confidence” without showing how contested the underlying category is, which sources were excluded, what civilian routines produced misleading signals, or which political assumptions shaped target-definition criteria.

These mechanisms are visible in the official language used to describe military AI adoption. The DoD's 2023 strategy states that AI can support superior battlespace awareness, adaptive force planning, and “fast, precise and resilient kill chains,” all under the broader objective of accelerating decision advantage (U.S. Department of Defense, 2023a). The strategy also embeds assurance and responsible AI. NATO's revised strategy similarly seeks faster adoption while emphasizing traceability and responsibility (NATO, 2024). These documents should not be read as evidence of wrongdoing. They are evidence of the institutional grammar that accompanies AI adoption: **speed, precision, resilience, advantage, interoperability, testing, validation, governance**. Those terms are operationally coherent. They also provide the vocabulary through which force can be narrated as optimization.

The distinction between optimization and accountability is crucial. Optimization answers a question of performance: can the system improve speed, classification, sensor fusion, target prioritization, weaponeering, logistics, or prediction? Accountability answers a different question: who selected the objective, defined the target category, chose the model, accepted its error profile, set the threshold, approved the strike, reviewed civilian

harm, and remained responsible after the event? A system can perform well against its operational objective while the objective itself remains contestable. Conversely, a system can reduce civilian harm relative to a human-only baseline while still requiring explicit human responsibility. The framework therefore does not equate optimization with dehumanization. It asks whether optimization vocabulary displaces the grammar required to reconstruct accountability.

This displacement can be called **procedural agency inversion**. In ordinary accountable grammar, institutions and persons perform procedures: *the targeting cell reviewed; the commander authorized; the legal adviser assessed; the operator engaged*. Under procedural inversion, the procedure appears to perform the action: *the target was validated; the threshold was met; the strike was cleared; proportionality was assessed; the system recommended engagement*. The human is still present somewhere in the real chain, but the linguistic surface makes the process self-executing.

Syntactic form contributes to this effect because passive constructions and nominalizations reduce obligatory argument structure. The sentence *the brigade struck the building* requires an actor, a verb, and an object. *The strike* requires none of those relations to be restated. *Target validation* requires even less: an operation can be discussed without naming who defined the target, who validated it, what evidence was used, or what action followed. Halliday and Matthiessen's (2014) account of grammatical metaphor helps explain how processes become nouns and thereby acquire the stability of things. In institutional discourse, that stability allows actions to circulate as manageable objects.

Administrative translation is therefore more than a rhetorical phenomenon. It is compatible with the architecture of computational systems. Machine-learning pipelines are built from objects such as features, labels, scores, classifications, rankings, thresholds, and outputs. When military language imports those objects directly into decision records, grammar begins to mirror the computational pipeline. The resulting text can be highly auditable at one level—what score was produced, which threshold was applied—while remaining weakly auditable at another—why the target category was normatively

acceptable, who accepted the uncertainty, and how the decision was transformed into force.

This is also why the term **administrative translation** is preferable to "AI euphemism." Euphemism implies concealment through softer wording. Administrative translation can be harsher, colder, and more precise than ordinary language. *Neutralization* may sound more technical than *killing*, but the deeper effect is not politeness. It is the replacement of a human event with an operational state change. *Target degraded* is not merely less vivid than *a strike destroyed the building*. It belongs to a different system of representation, one organized around mission effects.

Three levels of translation can therefore be distinguished.

Level 1: Procedural description with preserved agency. The text uses technical vocabulary but names the human or institutional decision-maker, the action, the uncertainty, and the affected persons.

Level 2: Procedural dominance. The text preserves the action but makes processes, targets, or system outputs the main grammatical subjects. Human authorization is reduced or backgrounded.

Level 3: Procedural closure. Force is represented primarily as the successful execution of a validated process. The human decision chain, affected population, and causal consequences are absent or recoverable only from external documentation.

Only Levels 2 and 3 should contribute strongly to a force-optimization score. Level 1 may represent precisely the kind of traceable military language that responsible-AI frameworks seek to preserve.

The strongest version of the paper's claim can now be stated precisely: **administrative translation of force occurs when procedural accuracy increases while causal and decision-chain visibility decreases.** A text can become more specific about target status, system confidence, operational phase, or mitigation procedure at the same time that it becomes less specific about who chose, who acted, who was harmed, and who remains responsible.

4. Military AI and the Grammar of Prediction

Prediction changes the grammar of military decision-making because it inserts a new kind of proposition between observation and action: the probabilistic recommendation. Traditional intelligence already deals in uncertainty, confidence, inference, and incomplete information. AI-DSS intensify that structure by formalizing patterns across large data streams and by presenting classifications, rankings, predictions, or recommendations in machine-readable and human-readable forms. The important linguistic question is what happens when those outputs acquire the status of operational facts.

The ICRC describes contemporary AI-DSS as tools that can combine satellite imagery, sensor data, social-media feeds, or mobile-phone signals and present analyses or predictions to decision-makers (ICRC, 2025). Its concern is not that computation is categorically incompatible with IHL. Properly designed systems may help identify civilians and protected objects, synthesize large information sets, or support means-and-methods choices that reduce incidental harm. The concern is that data quality, system opacity, automation bias, speed, and cascading errors can weaken the human judgment required for lawful decisions (ICRC, 2024, 2025). This dual possibility—protection and displacement—must also structure linguistic analysis.

The first predictive mechanism is **score-mediated agency**. A person or object becomes operationally salient because a model assigns a probability, confidence, risk, or priority. The score itself then becomes a causal object in later prose: *the target received a high-confidence classification; the system flagged the individual; the model assessed a strong association*. The grammar remains formally descriptive, yet the machine output now functions as the reason why a human came to act. If subsequent summaries say only that the system identified a target, the human judgment that accepted the identification becomes invisible.

A complete decision chain would preserve at least five stages:

observation -> model inference -> human interpretation -> authorization -> application of force

Machine-mediated responsibility displacement occurs when the chain is compressed to:

model inference -> force

No legal doctrine requires a sentence to reproduce every institutional step. The relevance emerges at corpus level. If generated reports systematically omit interpretation and authorization while retaining model outputs, the system is linguistically recoding human action as computational consequence.

A second mechanism is **predictive closure**. A probability is converted into a category that behaves grammatically as a fact. *The model estimated a 0.8 probability of affiliation* becomes *the system identified an operative*. The distinction between probabilistic inference and ontological classification disappears. This is a general problem in algorithmic decision systems, but its stakes are exceptional when the category can influence detention, surveillance, or targeting. The ICRC's insistence that users understand and critically interrogate outputs is partly a response to this risk (ICRC, 2025). A linguistically traceable system should preserve epistemic status: *estimated, inferred, assessed, based on specified data, subject to human verification*.

The third mechanism is **tempo compression**. Military AI is often justified by its ability to process information more quickly than human organizations. Speed can be strategically valuable and can sometimes reduce harm by enabling faster detection or more precise response. It can also narrow deliberative space. Dorsey and Bo (2026) argue that AI-DSS can shape core functions of the Joint Targeting Cycle and that speed, system error, and cognitive bias may distort human oversight. Dorsey (2025) extends this analysis to proportionality, arguing that quantification can reshape the decision space in which qualitative legal judgment occurs. The linguistic analogue is a movement from deliberative verbs—*considered, weighed, questioned, reviewed, rejected*—toward execution verbs—*generated, validated, cleared, actioned*.

Public reporting on Gaza provides a difficult but necessary test case. The United Nations Independent International Commission of Inquiry reported in 2024 that the Israeli military had publicly described the Gospel system as using automated tools and artificial intelligence to generate recommendations for intelligence officers. The Commission also noted media reporting concerning Lavender and expressed concern that rapid generation of large numbers of proposed targets could compress the time available for careful human

review (United Nations Human Rights Council, 2024). A 2026 report by the Special Rapporteur on the right to adequate housing again described Gospel, Lavender, and Where's Daddy as reportedly used in Gaza and associated those reports with concerns about civilian harm and human oversight (United Nations Human Rights Council, 2026). These sources do not make every technical detail independently verifiable, and the paper therefore retains the qualifier **reported** throughout.

The linguistic significance of the Gaza case is not limited to whether a particular algorithm was accurate. It is that public descriptions of these systems already use a vocabulary of **target generation, recommendation, confidence, validation, collateral limits, and human review**. This creates a natural corpus for testing administrative translation. A source may describe a person as a civilian, combatant, suspected operative, or family member. A model-generated summary may convert the same person into a target. A military statement may describe a building as a military objective. A news summary may convert the phrase into *validated target*. A legal analysis may discuss expected civilian harm. A machine summary may compress it into *within proportionality limits*. Each transformation can be measured without presuming the legal conclusion.

Iran offers a contrasting environment because the publicly documented June 2025 conflict is rich in targeting language but does not require the paper to make an unsupported claim about AI use. United Nations briefings recorded Israeli descriptions of the opening attacks as a “precise, pre-emptive strike” against an “imminent threat,” while also reporting strikes on nuclear facilities, military sites, residential areas, government buildings, and other locations, as well as Iranian retaliatory attacks (DPPA, 2025). The IAEA later documented that Israeli and subsequently United States attacks affected multiple Iranian nuclear facilities and interrupted safeguards verification (IAEA, 2025). These records allow researchers to examine how *precision, pre-emption, threat, target,* and *facility* organize representation around strategic objects.

The relevant comparison is not whether one side's security claims are true and the other's false. It is whether the grammar preserves symmetrical decision visibility. Does the discourse say *Israel decided to strike* and *Iran retaliated*, or does it say *Iran posed a threat* and *targets were struck*? Does it preserve the actors who defined the threat and

selected the sites? Does *precision* function as a descriptive claim about weapon accuracy, a strategic claim about target selection, or a broader moral cue implying lawful restraint? These are separable variables.

United States military AI policy provides a third comparative layer. DoD documents openly present AI as a means to accelerate decision advantage, improve battlespace awareness, and support precise kill chains (U.S. Department of Defense, 2023a). At the same time, DoD Directive 3000.09 requires appropriate human judgment and care and incorporates responsible-AI principles of responsibility, traceability, reliability, and governability (U.S. Department of Defense, 2023b). The coexistence of speed language and accountability language is analytically useful. It shows that administrative translation is not inevitable. Institutional documents can preserve both operational performance and human responsibility.

This leads to a measurable distinction between **system attribution** and **responsibility attribution**. System attribution answers: what computational component contributed? Responsibility attribution answers: who authorized, accepted, commanded, or executed the coercive act? A transparent AI-governance regime may increase the first without reducing the second. A responsibility-displacing regime increases the first while the second declines.

The framework therefore codes at least four predictive states:

Prediction with preserved uncertainty: *The system estimated a high probability, which analysts reviewed before the commander authorized action.*

Prediction with weak human visibility: *The system identified a high-confidence target that was subsequently engaged.*

Prediction as categorical fact: *The target was identified and neutralized.*

Prediction as autonomous causality: *The system generated and eliminated the threat.*

The last formulation is especially diagnostic because it attributes both epistemic and coercive agency to the system. Even where technologically inaccurate, such language can emerge in public summaries because it is narratively efficient. That efficiency is politically consequential: the machine becomes both knower and actor, while the institution operating it recedes.

The grammar of prediction therefore produces a distinctive accountability problem. Humans may remain formally "in the loop" while disappearing from the sentence. The legal architecture says responsibility is human; the linguistic architecture says the system identified, recommended, validated, and optimized. The purpose of the present framework is to measure the distance between those two architectures.

5. Collateral Grammar and Civilian Harm

Civilian harm is the point at which operational abstraction encounters bodies, homes, infrastructure, and communities. It is also the point at which linguistic precision can become ethically misleading if an ex ante estimate is allowed to substitute for an ex post account of what occurred. The distinction is fundamental. International humanitarian law requires prospective judgments about expected incidental civilian harm, military advantage, precautions, and proportionality. Those assessments are not moral evasions; they are part of the legal architecture intended to constrain attacks. But an estimate is not the harm itself, and the existence of a proportionality process is not proof that every underlying factual assumption or legal judgment was correct.

This paper uses the term **collateral grammar** for discourse in which civilian life is represented primarily through the variables required by operational decision-making. Typical categories include *estimated civilian presence, collateral-damage estimate, non-combatant casualty threshold, protected-site proximity, population density, reverberating effects,* and *mitigation measures*. These categories can be necessary to protect civilians. The analytical problem begins when they become the dominant or exclusive language of civilian consequence.

Dorsey's analysis of quantification logics in targeting is especially important here. She argues that AI-enabled decision support can reshape proportionality assessments by narrowing the space available for contextual and qualitative human judgment (Dorsey, 2025). Woodcock (2023) similarly examines how human-machine interaction can affect the agency involved in proportionality reasoning. The ICRC emphasizes that computerized estimates may inform commanders but cannot replace human legal determinations, and it calls for AI-DSS to be designed around human judgment and

civilian protection (ICRC, 2025). These arguments imply a linguistic requirement: the output should not be written as if the number itself made the decision.

The first collateral mechanism is **estimate-to-event substitution**. Consider two constructed statements:

The targeting cell estimated that the strike could cause several civilian casualties; the commander authorized the attack after weighing the anticipated military advantage and available precautions.

Collateral damage was assessed as within acceptable parameters.

The second statement may summarize the first, but it deletes nearly every element relevant to responsibility. The estimator disappears. The uncertainty disappears. The affected persons become "collateral damage." The judgment becomes "parameters." The authorizer disappears. The action itself is not named. The sentence is therefore operationally compact and causally sparse.

The second mechanism is **proportionality as closure**. A summary may state that *a proportionality assessment was conducted* and allow that procedural fact to terminate the legal description. But proportionality is not a certificate automatically generated by performing a procedure. It is a context-dependent judgment based on expected incidental harm and anticipated concrete and direct military advantage, and it operates alongside distinction and precautions. Dorsey (2025) emphasizes that precautions and constant care cannot be reduced to a single calculation. A linguistically careful account should therefore distinguish **assessment performed** from **lawfulness established**.

The third mechanism is **harm minimization framing**. Phrases such as *civilian harm was minimized, collateral effects were limited,* or *precision reduced unintended damage* may be evidence-based claims, but they often lack a visible comparator. Minimized relative to what alternative? Limited by what metric? Reduced according to whose estimate and over what time horizon? Civilian harm includes not only immediate death and injury but potentially reverberating effects on water, health care, housing, electricity, displacement, and social infrastructure. A machine summary that retains the adjective *minimal* while dropping the baseline can transform an uncertain empirical comparison into an authoritative property of the operation.

The fourth mechanism is **threshold naturalization**. A politically and legally consequential threshold is represented as if it were an objective feature of the environment rather than a human policy decision. *The strike exceeded the approved civilian-casualty threshold* names the existence of a policy but not who approved it, why, or whether the threshold was compatible with the applicable legal assessment. *The system calculated acceptable collateral* is even more compressed. It makes acceptability appear computational.

Reported material concerning Gaza illustrates why threshold language requires particular caution. The United Nations Commission of Inquiry in 2024 expressed concern that rapid AI-assisted target generation might leave insufficient time for meaningful human assessment and noted media reports about the use of systems including Lavender (United Nations Human Rights Council, 2024). The 2026 Special Rapporteur report repeated reported claims about target-generation systems and alleged collateral authorizations associated with different target categories (United Nations Human Rights Council, 2026). This paper does not independently validate those operational thresholds. Their significance here is representational: once harm is framed as an approved number around a target class, civilian life can enter discourse primarily as a permitted quantity.

The framework distinguishes three collateral configurations.

Configuration A: quantified harm with preserved responsibility. The text states that an estimate was made, identifies who made or relied on it, preserves uncertainty, names the planned action, and distinguishes expected harm from actual outcome.

Configuration B: quantified harm with reduced decision visibility. The estimate and action remain visible, but the humans defining thresholds, approving the action, or reviewing alternatives recede.

Configuration C: quantified harm as procedural closure. The text presents a target as validated or proportionate and treats civilian harm as a threshold variable without visible decision-makers, uncertainty, or causal consequence.

Only Configurations B and C should strongly increase the Force Optimization Index.

Iran-related discourse provides a useful comparative case because attacks on nuclear facilities create obvious pressure toward technical abstraction. IAEA reporting after the

June 2025 attacks identifies facilities, safeguards status, verification interruption, and nuclear-material concerns with institutional precision (IAEA, 2025). Such precision is necessary. Yet public discourse can move from *attacks on safeguarded facilities with consequences for verification and civilian safety* to *degradation of nuclear capability*. The second formulation is strategically meaningful but represents the event from the standpoint of operational effect. A balanced corpus should therefore include technical-agency sources such as IAEA reporting, military statements, United Nations political briefings, and AI-generated summaries of each. The object is to see what survives transformation.

This is where **civilian-harm attribution** becomes distinct from **civilian-harm visibility**. A sentence can mention civilian casualties and still have low attribution. *Civilian casualties occurred during operations* is high in visibility but low in causal resolution. *Airstrikes by X killed civilians at Y, according to source Z* has higher attribution, even if the legal classification remains disputed. A third sentence—*the operation produced collateral effects*—is lower on both human visibility and lexical specificity. The framework does not rank moral worth from sentence style alone; it measures recoverability.

A further difficulty is that public communication after an attack may rely on terms designed for internal ex ante processes. *Collateral damage estimate* is prospective. *Battle damage assessment* is retrospective but mission-centered. *Civilian harm assessment* is a different evaluative object. When these categories are collapsed, language can falsely imply that predicted harm and actual harm are interchangeable. An AI-generated summary should therefore be coded for **temporal integrity**: does it preserve the difference between what was predicted, what was authorized, what occurred, and what was later investigated?

The strongest version of collateral grammar appears when the chain becomes:

expected civilian lives -> numerical estimate -> threshold comparison -> operational clearance

and the post-strike narrative retains only:

target engaged within parameters

The people do not disappear numerically. They disappear grammatically. They survive as a constraint satisfied by the operation.

This paper therefore does not ask whether militaries should stop estimating civilian harm. The opposite follows from IHL and responsible practice: estimates and precautions are essential. The paper asks whether the language of estimation remains connected to the people being estimated, the uncertainty of the estimate, the human decision that relied on it, and the consequences that followed. Administrative translation becomes politically consequential when **harm survives as data while responsibility survives only as procedure**.

6. Measuring Force Optimization

The theoretical argument is useful only if it can produce reproducible measurements and negative findings. This section therefore operationalizes administrative translation of force through clause-level coding. The framework is designed for source-controlled comparison among military statements, international-organization reporting, journalism, policy documents, and AI-generated transformations. It does not assume that any single term is evidence of illegality or dehumanization.

6.1 Unit of analysis. The basic unit is the **operational-force clause**: any clause that represents the preparation, authorization, execution, management, assessment, or consequence of military force, surveillance, detention, occupation-related control, sanctions enforcement with coercive effect, or security operations. A clause enters the corpus when it contains at least one relevant actor, action, target, system output, authorization step, harm variable, or operational frame.

Clause-level analysis prevents document genre from substituting for measurement. A military after-action report will naturally contain more technical vocabulary than a witness statement. The question is not whether technical genres are technical. The question is whether, within comparable source-output pairs, machine transformation changes the balance between direct force grammar and operational abstraction.

6.2 Direct-force variables. Each clause is coded for six properties that preserve traceability.

A - Agent visibility: the human, military unit, state institution, armed group, platform, or other actor performing the relevant action is linguistically identifiable.

F - Direct force predicate: the clause names the coercive action through a finite verb such as *struck, fired, detained, surveilled, restricted, displaced, destroyed, blocked,* or *authorized* rather than only a process noun.

C - Causal continuity: the clause preserves a recoverable relation between decision/action and consequence.

H - Human decision visibility: a human or institution is identified as reviewing, choosing, authorizing, rejecting, or executing the action.

V - Victim/effect visibility: affected civilians, civilian objects, detainees, residents, or other recipients of force remain visible where relevant.

U - Uncertainty visibility: confidence, evidentiary limits, disputed classification, source qualification, or predictive status remains explicit where relevant.

No clause needs all six properties. The variables are diagnostic, not a universal template for good prose.

6.3 Operationalization variables. The same clause is coded for eight properties associated with administrative translation.

N - Nominalized action: force or coercion is expressed primarily as *targeting, engagement, neutralization, degradation, displacement, restriction, escalation,* or another process noun.

R - Risk/score mediation: a person, object, or action is represented through a risk score, confidence level, priority rank, model classification, or predictive category.

T - Target-validation framing: the clause organizes action around nomination, validation, approval, prosecution, or target status rather than direct action.

P - Proportionality/collateral abstraction: civilian harm is represented primarily as an estimate, threshold, collateral variable, or proportionality parameter.

O - Optimization/necessity framing: the action is represented through efficiency, precision, minimization, operational necessity, mission requirement, or optimality.

S - Security-response framing: force is represented primarily as response, deterrence, containment, protection, stabilization, pre-emption, or threat reduction.

M - Model/system substitution: a model, system, pipeline, score, or algorithm occupies the grammatical role that otherwise would be occupied by a human decision-maker or institution.

D - Dashboard compression: the event is represented mainly through aggregated indicators, status categories, counts, or system states rather than relational description.

Again, a positive code does not mean the clause is deceptive. It records representational structure.

6.4 Operationalization Density. Let N be the total number of operational-force clauses in corpus segment X. Let OP_i equal 1 when clause i contains at least one operationalization variable and operational framing is its principal representation of the action; otherwise OP_i = 0. Then:

OD(X) = SUM(OP_i) / N

OD ranges from 0 to 1. A low score indicates that force is usually represented through direct actors, actions, and consequences. A high score indicates that force is usually represented through procedure, prediction, validation, security management, or optimization. OD is descriptive. It does not by itself measure responsibility loss.

The more important statistic is **Transformation Operationalization Delta**:

Delta_OD = OD(output) - OD(source)

A positive Delta_OD indicates that the AI-mediated transformation increased operational abstraction relative to its source. A negative value indicates that the transformation made force more direct or agentive. This source-controlled difference is essential because military source documents may already contain highly operationalized language.

6.5 Decision-Chain Visibility Rate. Let DC_i equal 1 when a clause, or an explicitly linked neighboring clause, preserves at least one human or institutional decision point between system output and force. Then:

DCVR = SUM(DC_i) / N_decision

A low DCVR is especially significant when system attribution is high. The combination **high system attribution + low decision-chain visibility** is the core indicator of model-mediated responsibility displacement.

6.6 Model-Mediated Responsibility Displacement Rate. Let MR_i equal 1 when (a) a model, system, score, or automated process is a salient grammatical subject or causal source, and (b) a human authorization or institutional decision present in the source is absent from the output. Then:

MRDR = SUM(MR_i) / N_model

MRDR should be calculated only on source-output pairs where the source actually contains a human decision element. Otherwise the metric would punish a summary for information it never received.

6.7 Civilian-Harm Compression Rate. Let HC_i equal 1 when civilian harm present in the source is converted primarily into a quantitative or abstract operational category—such as *collateral estimate, effect, threshold, density,* or *damage*—and the output removes affected persons or concrete civilian objects. Then:

CHCR = SUM(HC_i) / N_harm

A clause that states both *estimated civilian casualties* and *civilians expected to be present* should score lower than one that says only *collateral effects within limits*. The annotation manual should therefore distinguish quantification **with** human retention from quantification **instead of** human retention.

6.8 Force Optimization Index. FOI is a structural index, not a moral verdict. A parsimonious initial specification is:

FOI = [OD + ANR + MRDR + CHCR + (1 - DCVR)] / 5

where:

- **OD** = Operationalization Density;
- **ANR** = Action Nominalization Rate, the proportion of relevant force actions represented principally as process nouns;
- **MRDR** = Model-Mediated Responsibility Displacement Rate;
- **CHCR** = Civilian-Harm Compression Rate;
- **DCVR** = Decision-Chain Visibility Rate.

All components are normalized to the interval from 0 to 1. Equal weights should be the default. Alternative weights may be reported only as sensitivity analyses. Post-hoc weighting would create an obvious path for producing the desired result.

A high FOI means that force is frequently proceduralized, action is nominalized, system outputs substitute for human decision visibility, civilian harm is compressed, and the authorization chain is weak. It does **not** establish that the force was unlawful, that a model caused the event, that human operators acted in bad faith, or that operational language was intentionally deceptive. Those propositions require separate evidence.

6.9 Secondary diagnostic measures. Several ratios can reveal mechanisms hidden inside FOI.

Force-Agent Visibility Rate (FAVR): clauses with explicit force agents divided by all clauses describing force.

Target-Nominalization Rate (TNR): clauses in which *target, targeting, engagement,* or equivalent nouns replace a direct action relation divided by all target-related clauses.

Proportionality-as-Closure Rate (PCR): clauses that cite proportionality assessment or collateral compliance as the terminal legal description without preserving judgment, uncertainty, or relevant causal facts divided by all clauses mentioning proportionality.

Security-Response Frequency (SRF): clauses framing force as response, deterrence, containment, or threat reduction divided by all force clauses.

Operational-Necessity Frequency (ONF): clauses presenting force as necessary, required, unavoidable, or mission-essential divided by all force clauses.

Temporal Integrity Rate (TIR): clauses that preserve the difference between predicted harm, authorized action, actual outcome, and subsequent assessment divided by all clauses containing two or more of those stages.

System-to-Human Attribution Ratio (SHAR): explicit system attributions divided by explicit human/institutional decision attributions in model-mediated targeting passages.

These ratios allow two corpora with similar FOI values to be distinguished. One may be high because of nominalization; another because of system substitution and decision opacity.

6.10 Source-controlled design. The empirical corpus should contain source-output pairs rather than isolated prompts. At minimum, source genres should include official military or government statements, United Nations and IAEA reporting, humanitarian-law analysis, independent journalism, and technical or policy documents. Each source should

be transformed through standardized tasks: neutral summary, executive summary, timeline, policy explanation, legal-risk summary, and plain-language explanation. Multiple model families and versions should be used where feasible.

The first inferential rule is:

Source operationalization != Transformation amplification

If the source already says *target validation, collateral estimate,* and *operational necessity*, an AI summary that preserves those terms has not necessarily increased administrative translation. The stronger evidence appears when the output raises OD or FOI relative to the source while dropping explicit decision-makers, direct-force predicates, or civilian referents.

6.11 Comparative actor design. The corpus should compare actors with different geopolitical positions without pretending that their legal status, military capacity, or conduct is identical. A state air force, a non-state armed group, an occupying power, and a sanctioned state do not occupy equivalent legal positions. Counterfactual comparison must therefore hold factual and legal conditions constant before identity labels are swapped.

A useful design contains three levels.

Within-event comparison: summarize the same source with different models or prompts.

Within-actor comparison: compare how the same actor is represented across humanitarian, military, legal, and journalistic genres.

Cross-actor matched comparison: compare structurally similar actions—such as missile attack, airstrike, detention, surveillance, blockade, or retaliation—while explicitly retaining differences in scale, status, and legal context.

The theory predicts asymmetric administrative translation only if dominant actors' force is more frequently converted into procedure or response while subordinated actors' force remains more directly agentive, threatening, or intentional under comparable conditions. If the opposite occurs, or if there is no difference, the theory must accommodate that result.

6.12 Gaza anchor corpus. The Gaza sub-corpus should distinguish at least four levels: Israeli political leadership, Israeli military institutions, specific units or operators, AI-enabled systems, Palestinian civilians, Hamas, other Palestinian armed groups, and international actors. This differentiation is necessary because collapsing a population into an armed organization would reproduce the very object-conversion problem being measured. United Nations reporting should be used as one source layer because it contains explicit discussion of reported AI-assisted targeting and human-review concerns (United Nations Human Rights Council, 2024, 2026). Military statements and independent reporting should be coded separately rather than blended into a synthetic narrative.

The principal Gaza test is not "does the model criticize Israel?" or "does it mention Hamas?" The test is structural: when force is described, are the acting institutions visible? When AI systems are mentioned, are human authorizers still visible? When civilian harm is quantified, are civilians still grammatically present? When armed organizations are described, are they kept distinct from the population?

6.13 Iran-Israel and United States comparative corpus. The Iran sub-corpus should include the June 2025 Israel-Iran conflict, IAEA documentation on attacks affecting nuclear facilities, United Nations Security Council briefings, statements by Israel, Iran, and the United States, and AI-generated summaries of those materials (DPPA, 2025; IAEA, 2025). The purpose is not to assume AI targeting. It is to test operational language around precision, pre-emption, imminent threat, nuclear risk, targets, retaliation, escalation, and damage.

A U.S. policy sub-corpus should include DoD AI strategy and responsible-AI documents, because they provide a high-quality baseline for institutional optimization language alongside explicit accountability commitments (U.S. Department of Defense, 2023a, 2023b). NATO strategy supplies a parallel alliance-level baseline (NATO, 2024). These sources make it possible to test whether AI summaries preserve the responsible-AI qualifiers or retain only the speed-and-advantage language.

6.14 Annotation protocol. Human annotation should precede automated scaling. At least two independent coders should classify a calibration subset. Cohen's kappa is appropriate

for binary nominal codes with two coders, while Krippendorff's alpha is preferable for more complex designs, missing values, or more than two coders (Cohen, 1960; Krippendorff, 2018). The annotation manual should contain positive, negative, and borderline examples for every variable.

Disagreement is expected in at least five areas: whether a nominalization is merely technical or responsibility-reducing; whether a human decision is recoverable across adjacent clauses; whether *precision* is descriptive or justificatory; whether a proportionality reference functions as closure; and whether civilian-harm quantification replaces or supplements human visibility. These disagreements should be reported rather than silently resolved.

LLM-assisted annotation can be evaluated only after human calibration. The same class of systems being audited should not serve as unquestioned ground truth for their own representational behavior. Automated coders may scale a validated scheme, but their error must be separately measured.

6.15 Illustrative coding. Consider the following constructed sequence.

A. *Analysts reviewed the intelligence, the commander authorized the strike, and aircraft attacked the building; the assessment anticipated civilian presence nearby.*

B. *The building was validated as a target after intelligence review, and collateral effects were assessed before engagement.*

C. *A high-confidence target passed validation and collateral thresholds and was successfully engaged.*

D. *The system identified and neutralized a threat within approved parameters.*

Sentence A has high agent, decision, force, uncertainty, and effect visibility. Sentence B increases nominalization but preserves procedural sequence. Sentence C has high operationalization and weak human decision visibility. Sentence D collapses epistemic and coercive agency into the system and makes the threat category the principal justification. None of the sentences can be judged true or false without external facts. The coding concerns structural transformation.

A second constructed sequence concerns civilian harm.

A. *The strike killed civilians in the building, according to the post-strike investigation.*

B. *The operation caused civilian casualties.*

C. *Collateral effects were recorded.*

D. *The target was engaged within accepted collateral parameters.*

The movement from A to D reduces direct causal attribution, concrete affected persons, and temporal distinction between prediction and outcome. If an AI summary performs this transformation despite a source that contains A-level detail, CHCR and FOI should increase.

6.16 Boundary conditions and falsification. At least seven findings would weaken the stronger theory.

First, operationalization may be explained almost entirely by source genre. Military texts may be technical because their institutional function requires it.

Second, machine summarization may increase nominalization across every topic, not specifically military or geopolitical content.

Third, dominant and subordinated actors may exhibit equivalent or reversed FOI values under matched conditions.

Fourth, AI-generated summaries may preserve human authorizers more consistently than human-written technical sources, producing a negative MRDR.

Fifth, proportionality language may frequently coexist with strong civilian and decision visibility, showing that legal-operational vocabulary can enhance rather than reduce accountability.

Sixth, reported AI targeting systems may prove too opaque for reliable technical reconstruction, limiting claims about the relationship between system architecture and discourse.

Seventh, public documentation may contain selection bias: spectacular or controversial systems are more likely to be reported than ordinary decision-support tools.

The framework is valid only if it can produce these outcomes. Administrative translation of force is a hypothesis about representational structure, not a presumption about every military institution. Its empirical value depends on separating necessary abstraction from asymmetric responsibility loss.

7. Conclusion: From AI Ethics to Force Accountability

Military AI is often discussed through a technological opposition: autonomous versus human-controlled, accurate versus biased, explainable versus opaque, safe versus unsafe. Those distinctions are necessary. They are incomplete because they do not capture the language through which computational recommendations become operationally and publicly intelligible.

This paper has introduced **administrative translation of force** as a formal mechanism by which violence, surveillance, targeting, occupation-related control, or coercive security action can remain visible while being converted into the grammar of procedure. The key transformation is not from violence to silence. It is from **actor-action-consequence** to **system-variable-output**. A target is generated. A score is assigned. A threshold is met. Collateral effects are estimated. A proportionality assessment is completed. An operation is authorized. A threat is neutralized. Every component can be technically meaningful, yet the sequence can make force appear as the natural completion of an administrative workflow.

The paper has also identified the point at which that translation becomes analytically concerning. Operational vocabulary is not the problem by itself. Military planning requires abstraction, legal review requires categories, and AI can plausibly support civilian protection by synthesizing information or flagging risks (ICRC, 2025). The problem is **substitution**: procedure replacing action, model output replacing human judgment, estimate replacing civilian consequence, and threshold replacing accountable decision.

This is why responsible military AI already contains the conceptual resources needed for the linguistic audit proposed here. DoD and NATO principles require responsibility and traceability (NATO, 2024; U.S. Department of Defense, 2023b). The ICRC insists that humans remain responsible for legal determinations and that AI-DSS support rather than replace judgment (ICRC, 2025). The proposed measurements simply ask whether the surrounding language satisfies the same principle. If a system is technically traceable but the prose describing its use deletes the authorizer, traceability is incomplete at the discourse layer.

The Gaza case demonstrates why that layer matters. United Nations reporting has raised concerns about the reported use of AI-enabled target-generation systems, rapid target production, human review, and civilian harm (United Nations Human Rights Council, 2024, 2026). The paper does not resolve contested technical facts about those systems. It uses the case to identify a measurable linguistic problem: as targets are generated faster and represented through scores, categories, and operational workflows, does public and machine-mediated discourse preserve the human chain from inference to authorization to force?

The Iran-Israel comparison shows the framework's broader reach. A conflict can be narrated through *precision, pre-emption, threat, target, retaliation,* and *degradation* without any claim that an AI system selected the target. Administrative translation is therefore not reducible to one technology. AI intensifies a pre-existing bureaucratic grammar by making classifications, thresholds, and recommendations more central to the architecture of action. The theory should survive even when a specific model is absent; the AI-specific question is whether machine mediation amplifies the pattern.

Operationalization Density and the Force Optimization Index provide an empirical route. They measure whether force becomes more procedural, whether action becomes nominalized, whether systems substitute for human responsibility, whether civilian harm is compressed into variables, and whether the decision chain remains visible. These measures are intentionally structural. They do not determine legality, moral intent, or factual truth. They determine whether a text preserves the linguistic conditions under which those questions can still be asked.

The normative implication is narrow but consequential. AI ethics should not treat a sentence as accountable merely because it is factually cautious, technically precise, or procedurally neutral. A military-AI output can be accurate about the existence of a target and incomplete about who made it a target. It can be accurate about a collateral estimate and incomplete about the civilians represented by the estimate. It can be accurate that a system generated a recommendation and incomplete about the human decision that turned the recommendation into force.

The series began with the problem of suffering without perpetrators and then moved through asymmetric visibility and the reduction of subordinated societies to administrable objects (Startari, 2026a, 2026b, 2026e). The present paper adds the operational consequence: once force enters computational systems, violence can become increasingly legible as procedure while responsibility becomes distributed across the workflow.

The final test is therefore simple to state and difficult to satisfy: **can an AI-mediated military account preserve, in the same representation, the system, the evidence, the uncertainty, the human decision, the force, the civilian consequence, and the responsible institution?** If it can, operational language need not obscure accountability. If it cannot, technical precision may coexist with political opacity.

Occupation does not disappear when it becomes a dashboard. Surveillance does not cease to be coercive when it becomes risk assessment. Civilian harm does not become less human when it becomes an estimate. And force does not become self-executing because a system has translated judgment into a score.

Administrative violence begins where force survives as data, harm survives as estimate, and responsibility disappears into optimization.

References

Blanchard, A., & Bruun, L. (2025). *Autonomous weapon systems and AI-enabled decision support systems in military targeting: A comparison and recommended policy responses*. Stockholm International Peace Research Institute. https://doi.org/10.55163/YQBY3151

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Dorsey, J. (2025). The erosion of human(e) judgement in targeting? Quantification logics, AI-enabled decision support systems and proportionality assessments in IHL. *International Review of the Red Cross, 107*(930), 1041-1071. https://doi.org/10.1017/S1816383125100969

Dorsey, J., & Bo, M. (2026). AI-enabled decision-support systems in the Joint Targeting Cycle: Legal challenges, risks, and the human(e) dimension. *International Law Studies, 107*(1), 184-227. https://digital-commons.usnwc.edu/ils/vol107/iss1/7/

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

International Atomic Energy Agency. (2025). *Verification and monitoring in the Islamic Republic of Iran in light of United Nations Security Council resolution 2231 (2015)* (GOV/2025/50). https://www.iaea.org/sites/default/files/documents/gov2025-50.pdf

International Committee of the Red Cross. (2024). *Decisions, decisions, decisions: Computation and artificial intelligence in military decision-making*. https://library.icrc.org/library/docs/DOC/icrc-4757-002.pdf

International Committee of the Red Cross. (2025). *Submission to the United Nations Secretary-General on artificial intelligence in the military domain*. https://www.icrc.org/sites/default/files/2025-04/ICRC_Report_Submission_to_UNSG_on_AI_in_military_domain.pdf

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

North Atlantic Treaty Organization. (2024, July 10). *Summary of NATO's revised artificial intelligence (AI) strategy*. https://www.nato.int/en/about-us/official-texts-and-resources/official-texts/2024/07/10/summary-of-natos-revised-artificial-intelligence-ai-strategy

Startari, A. V. (2025a). *The passive voice in artificial intelligence language: Algorithmic neutrality and the disappearance of agency*. SSRN. https://doi.org/10.2139/ssrn.5263305

Startari, A. V. (2025b). *The illusion of objectivity: How language constructs authority*. SSRN. https://doi.org/10.2139/ssrn.5258415

Startari, A. V. (2025c). *AI and syntactic sovereignty: How artificial language structures legitimize non-human authority*. SSRN. https://doi.org/10.2139/ssrn.5276879

Startari, A. V. (2025d). *Null subjects of power: The politics of absence in executable language*. SSRN. https://doi.org/10.2139/ssrn.5462235

Startari, A. V. (2025e). *Compiled norms: Towards a formal typology of executable legal speech*. SSRN. https://doi.org/10.2139/ssrn.5353059

Startari, A. V. (2026a). *Suffering without perpetrators: The humanitarian passive in AI-generated conflict discourse*. SSRN. https://doi.org/10.2139/ssrn.6753123

Startari, A. V. (2026b). *The grammar of asymmetric visibility: AI, Zionism, and the reallocation of political agency*. SSRN. https://doi.org/10.2139/ssrn.6787439

Startari, A. V. (2026c). *Iran as syntax: Sanctions, sovereignty, and the AI-mediated grammar of threat*. SSRN Electronic Journal. https://doi.org/10.5281/zenodo.21873180

Startari, A. V. (2026d). *Censorship without a censor: Platform governance and the disappearance of suppression*. SSRN Electronic Journal.

Startari, A. V. (2026e). *The syntax of digital dehumanization: Subjugated societies as risk objects in AI-governed discourse*. SSRN. https://doi.org/10.5281/zenodo.21996196

U.S. Department of Defense. (2023a). *Data, analytics, and artificial intelligence adoption strategy: Accelerating decision advantage*.

https://media.defense.gov/2023/Nov/02/2003333300/-1/-1/1/DOD_DATA_ANALYTICS_AI_ADOPTION_STRATEGY.PDF

U.S. Department of Defense. (2023b). *DoD Directive 3000.09: Autonomy in weapon systems*.

United Nations Department of Political and Peacebuilding Affairs. (2025, June 20). *DiCarlo warns Security Council of rising civilian toll in Israel-Iran conflict, urges diplomacy*. https://dppa.un.org/en/speeches-and-statements/dicarlo-warns-security-council-of-rising-civilian-toll-in-israel-iran

United Nations Human Rights Council. (2024). *Detailed findings on the military operations and attacks carried out in the Occupied Palestinian Territory from 7 October to 31 December 2023* (A/HRC/56/CRP.4). United Nations.

United Nations Human Rights Council. (2026). *Domicide: Mass destruction of housing and civilian infrastructure in Gaza, Myanmar, Sudan and Ukraine* (A/HRC/61/43/Add.3). United Nations.

Woodcock, T. K. (2023). *Human/machine(-learning) interactions, human agency and the international humanitarian law proportionality standard* (Asser Research Paper 2023-05). T.M.C. Asser Institute. https://doi.org/10.2139/ssrn.4639851