

A Less Dangerous Use of AI in Armed Conflict?

Decision Support Systems under the Lens of International Law

Daniele Amoroso

Università degli studi di Cagliari



+972
MAGAZINE

'Lavender': The AI machine directing Israel's bombing spree in Gaza



Outline

- The basics of targeting law
- The (IDF's) targeting process
- The +972 report on Lavender
- The systemic risks of heavily embedding AI in the targeting process
 - Weakened human oversight and the principle of precaution
 - Oversimplification of legal rules: the trivialization of proportionality
- Conclusions

The basics of targeting law

Principle of distinction

- Attacks cannot be directed against civilians and civilian objects

Principle of proportionality

- For each attack, the expected incidental damage to civilians and civilian objects cannot be excessive compared to the anticipated military advantage

Principle of precaution

- Everything feasible must be done to ensure that attacks are directed against legitimate targets and that civilian harm is avoided/minimized

The IDF's Air Force targeting process

Target Development

Target Assessment

Pre-Strike Controls

Strike Operations

The Role of the IDF's International Law
Department

Highlights from the +972 report

Part I: Lavender, the target machine

+972
MAGAZINE

- All operatives of Hamas military wing considered as «human targets»
Challenge: how to identify them? => The AI-based Lavender system
- How Lavender (allegedly) worked: red flags
 - Fed with huge amount of data from Gaza population => mass surveillance
 - Did assign an «affiliation to Hamas» score based (also) on ambiguous indicators (eg being in a WhatsApp group with a known Hamas militant)
 - Trained also with mislabeled data (eg inclusion in the data set of the profiles of Gaza civil defence workers)

10% false positive rate => human validation?

Highlights from the +972 report

Part II: The «Lavender protocol»

+972
MAGAZINE

- Validation by human operators
 - Lavender's targeting recommendations treated as «superior orders»
 - Human override based on gender check (no women in Hamas military wing)
- Standard targeting plan: strike the targets when they are at home
 - «Where's Daddy» tracking system
 - In many cases, no pre-strike check
- Proportionality threshold pre-approved by the Int. Law Dep.
 - Impossible case-by-case assessment
 - High threshold: 15-20 per junior militant
- Use of «dumb bombs»
 - Laser-guided munitions not to be wasted on low value militants
 - Increase of incidental civilian casualties

Israel Defence Forces' response to claims about use of 'Lavender' AI database in Gaza

Statement issued after six Israeli intelligence sources speak out about AI use in Gaza war

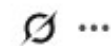


The “system” your questions refer to is not a system, but simply a database whose purpose is to cross-reference intelligence sources, in order to produce up-to-date layers of information on the military operatives of terrorist organizations. This is not a list of confirmed military operatives eligible to attack.

For each target, IDF procedures require conducting an individual assessment of the anticipated military advantage and collateral damage expected. Such



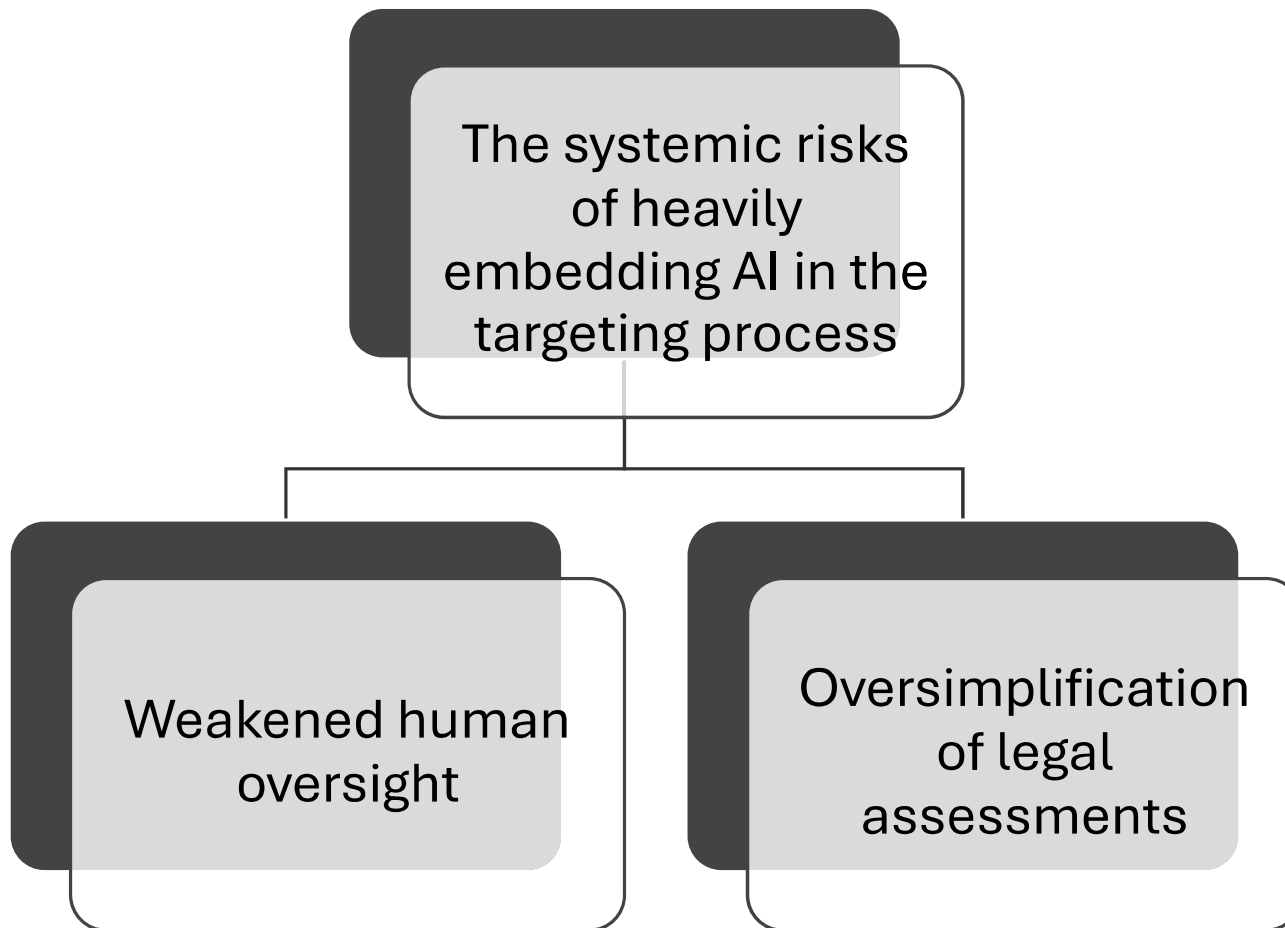
Israeli Air Force 
@IAFsite



The airstrikes killed hundreds of terrorists and attacked over 3,600 targets, including: command and control targets, strategic military infrastructure, weapons production sites, intelligence assets, leadership targets, naval superiority targets and rocket system targets.

5:03 pm · 12 Oct 2023 · **457.3K** Views

What the Lavender case teaches us



Weakened human oversight and the principle of precaution

- The principle of precaution
 - Positive duty instrumental to distinction and proportionality
 - Everything feasible must be done to ensure correct target identification
- Implications for «human-machine» interactions
 - AI can make serious mistakes that human double-checking would easily catch
 - Human «fail-safe» as a legal requirement
- Factors eroding human oversight
 - Automation bias
 - WYSIATI heuristics (AI target lists «anchor» decisions)
 - Authority bias (Lavender's outputs equated to «superior orders»)
 - Time pressure and institutional disincentives to override AI

Oversimplification of legal rules: The trivialization of proportionality

- The principle of proportionality
 - The «core»: the assessment of excessiveness
 - The normative standard of the «reasonable military commander»
 - ⇒ Hard to «code» into an algorithm
- A suggested workaround: pre-determined casualty threshold
 - Int. Law. Dep. standardized proportionality checks in the Lavender case
- Legal issues
 - Individualized proportionality assessments turned into a blanket permissive rule
 - Not unlawful per se, but it makes it more likely disproportionate strikes (unless the threshold is set very low)
- Ethical issue
 - «Normalization» of civilian deaths (box-ticking exercise)



Conclusions

- The need for Meaningful Human Control over AI-DSS
- What makes human control over AI-DSS «meaningful»?
 - Human-Centered Design
 - Clear Distribution of Human Responsibilities
 - Reflective Training
- The way forward
 - Need for new law vs. crisis of multilateralism
 - Embedding MHC pillars in national doctrines and military manuals