

AI Warfare and the Law

**by
Bill Boothby**


All rights associated with this presentation are reserved by and to William Boothby




Introduction

- ◆ Alan Turing 1950 article – Can computers think?
- ◆ Motivation – operationalising law on AI warfare + cdr not always responsible

Narrow & General AI

- ◆ Computers use patterns in data to create a data model to make predictions
 - ◆ General AI uses deep learning to perform theoretical tasks, develops neural networks, apply judgment and reasoning, evaluate possibilities, identify and classify phenomena
- 

Legally important aspects:

- ◆ Black box
 - ◆ Both Autonomy and decision support
 - ◆ Generative AI – e.g. CHAT-GPT - Deep fake
 - ◆ BUT
 - ◆ AI may misperform – production level, hacking, error, bias, sychophancy, deception etc ... and operator may not know
- 


AI and UN Charter Law

- ◆ Faulty ad bellum decisions dangerous
 - AI can misinterpret its task
 - Are all causes of distorted output excluded?
 - Is cybersecurity robust and maintained?
 - Do planners know AI vulnerabilities?
 - Does system disclose when not operating well?


Weapons Law

- ◆ Consider rapid threats
- ◆ No ad hoc law on AI
- ◆ Will AI perform as intended – Testing – ‘it can’... vs ‘it will’... - realistic testing
- ◆ Don't equip with illegal warheads
- ◆ Can targeting law be applied e.g. by operator?

Applying distinction to persons

- ◆ Will AI distinguish combatants - civilians – hors de combat?
 - ◆ Does AI recognise doubt?
 - ◆ Will AI apply doubt rule?
 - ◆ Are biometrics reliable?
- 

Distinction and Objects

- ◆ What was 'object of attack' – black box
 - ◆ Primary purpose to spread terror – can AI have a purpose?
 - ◆ Will AI reliably I/D mil obj by nature? Checked by testing? Is reasoning transparent?
 - ◆ Location – will AI assess operational context - Human pre-mission briefing of AI might help
 - ◆ Purpose – will AI assess future enemy use?
 - ◆ Book breaks down rules into elements
- 


AI in hybrid warfare

- ◆ A vehicle for action in enemy State, e.g.:
- ◆ key infrastructure disruption;
- ◆ propaganda broadcasting;
- ◆ false reports suggesting war crimes;
- ◆ Misrepresenting what is going on


International criminal law

- ◆ Machines do not have criminal responsibility
- ◆ Intent and knowledge required
- ◆ Cdr unaware of crime lacks intent/knowledge
- ◆ Who is 'perpetrator'? – Software designer?
- ◆ **Command responsibility - Should commander have known?**
- ◆ **Who 'directs' autonomous attack where General AI selects target and arranges attack?**


Some Neutrality Law implications

- ◆ Will AI comply with neutrality?
 - ◆ Must not fire in, to, from or through neutral territory
 - ◆ If neutral aware, must do what it can to end violation
 - ◆ If belligerent aware, must end violation
 - ◆ Implies transparency and no falsification by AI of location
- 


CCW LAWS discussions

- ◆ Autonomy is focus
 - ◆ Ongoing since 2014
 - ◆ Now discussing text options
 - ◆ Differing perspectives among States
 - ◆ 11 Guiding Principles
 - ◆ Maybe a 2-tier approach – prohibitions and restrictions
 - ◆ Outcome uncertain
- 
- A stylized teal silhouette of a mountain range is located in the bottom right corner of the slide, partially overlapping the text area.


Other law

- ◆ Human rights law (applies in armed conflict – jurisdiction?)
 - ◆ Applicable domestic law (where AI not used in hostilities)
 - ◆ Contract law
 - ◆ Tort of negligence?
 - ◆ Product liability?
 - ◆ Employment law
 - ◆ Service discipline
 - ◆ Inquiries where things go wrong
- 
- A stylized teal silhouette of a mountain range is located in the bottom right corner of the slide, partially overlapping the text area.

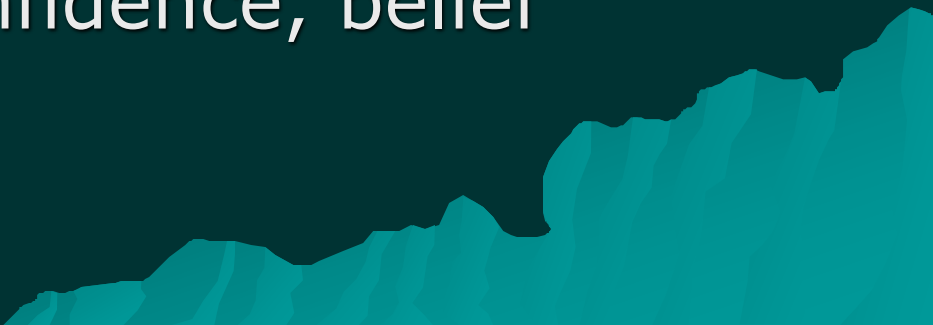
Responsibility

- Did an act lie within a person's duties?
 - Did s(he) fail to do it to required standard?
 - Did bad outcome result?
 - Should s(he) be blamed?
 - Are other factors jointly responsible?
 - What was primary cause?
- 

Human roles in AI

- ◆ System designers
 - ◆ Manufacturers & suppliers
 - ◆ Procurement processes
 - ◆ Specifiers of usage
 - ◆ Weapon testing
 - ◆ Legal review
 - ◆ Acquisition
 - ◆ Information transmission to users
 - ◆ Not just Commanders & operators
 - ◆ I.e. the supply chain
- 
- A stylized teal silhouette of a mountain range is located in the bottom right corner of the slide, partially overlapping the dark teal background.

A final word

- ◆ Anthropomorphising – ‘determine’
 - ◆ The meaning of words an ongoing challenge in this project
 - ◆ Aim of book to operationalise application of law to AI warfare
 - ◆ An iterative process
 - ◆ AI still work in progress – virtual assistants!
 - ◆ Meaning of think, confidence, belief
- 

The book

- ◆ AI Warfare and the Law
- ◆ Bill Boothby
- ◆ <https://digital-commons.usnwc.edu/ils/vol104/iss1/1/>.

Questions?

