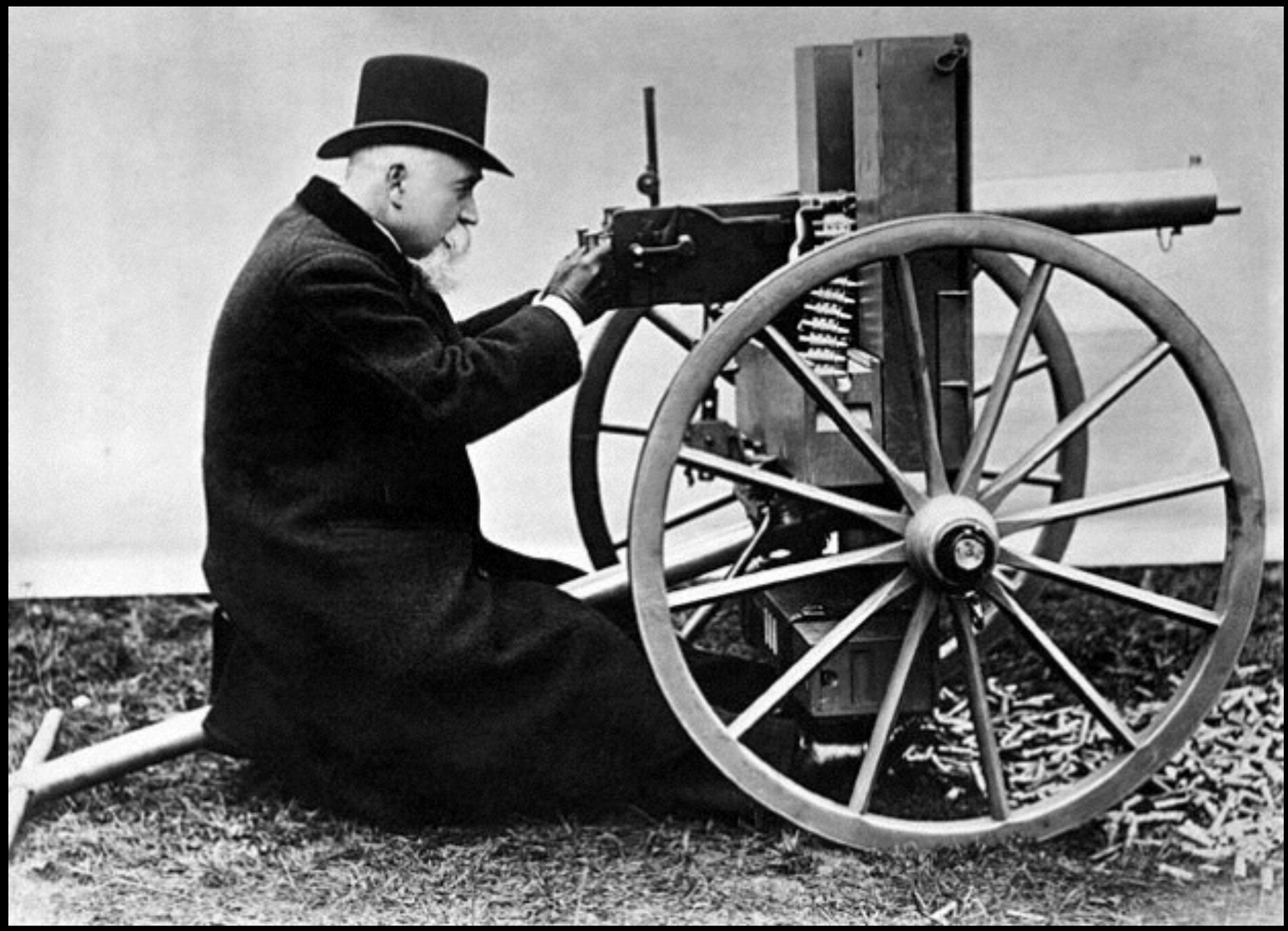
DALI 2016 - Data Learning and Inference

Machine Learning and Society

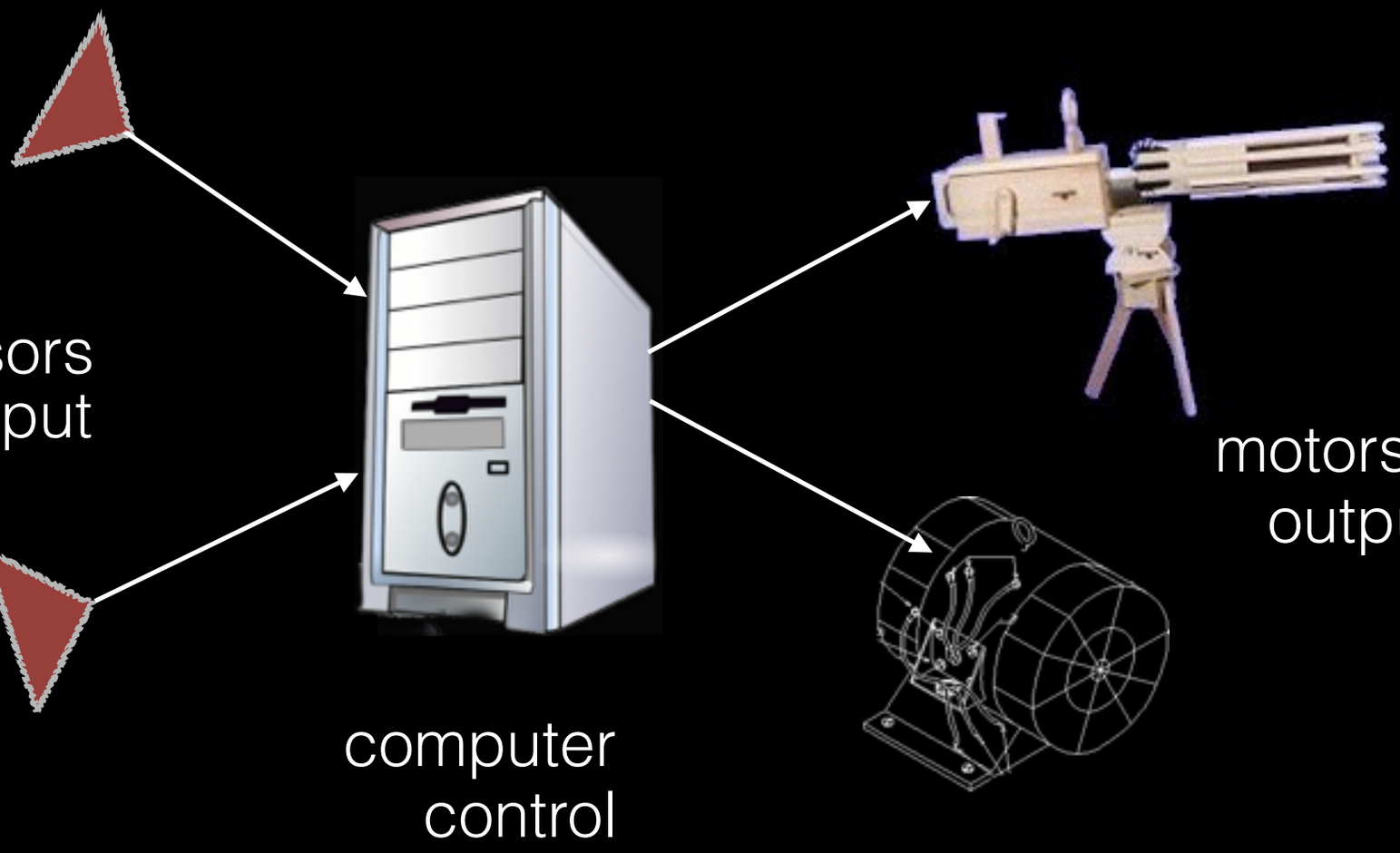
Why Autonomous Warfare is a Bad Idea

Noel Sharkey
University of Sheffield
International Committee for Robot Arms Control
Foundation for Responsible Robotics

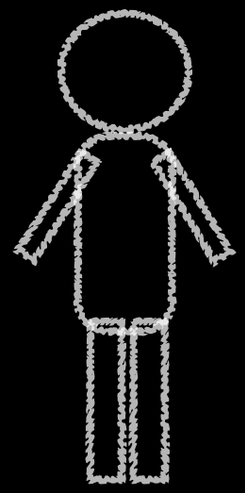
direct human control of weapons



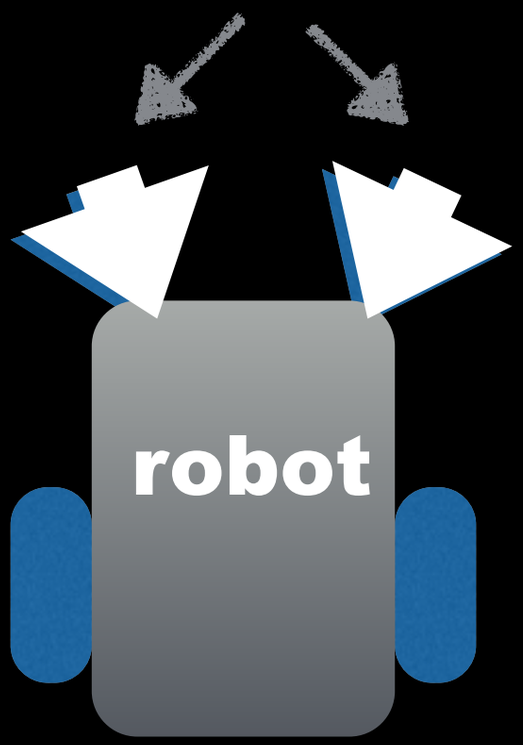
autonomous weapons control



Animation showing a simple version of the kill decision. It is static in this pdf.



heat sensors



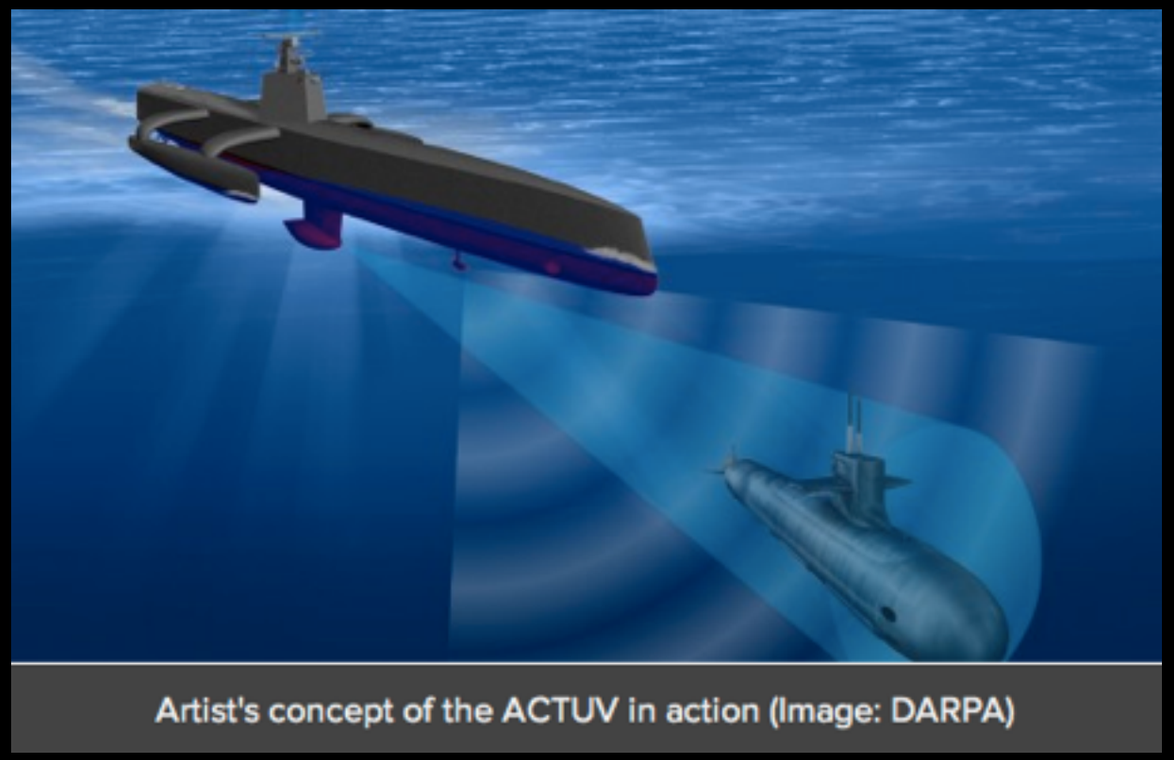
PROGRAM

if heat detected on one sensor
    rotate robot
        until both sensors detect heat
then
    fire weapons

US: autonomous X47-b

UK: Taranis autonomous intercontinental combat aircraft

Israel: autonomous Guardium



US: autonomous submarine hunting sub

US: CRUSHER

China: Anjian air to air combat

SWARM

THE FUTURE IS NOW

4 major problem areas

I. over reliance on computer programs

II. compliance with IHL

III. ethical compliance

IV. impact on gobal security

I. Possible failures (DoD 2012)

human error,

human-machine interaction failures,

malfunctions,

communications degradation,

software coding errors,

enemy cyber attacks

infiltration into the industrial supply chain,

jamming, spoofing, decoys,

other enemy countermeasures or actions, unanticipated situations on the battlefield

International Humanitarian Law (IHL)

II. Compliance with international humanitarian law?

★ Principle of distinction

★ Principle of proportionality

★ Precaution

★ Accountability


EXIT
Open

Autonomous Harpy radar killer

Made by IAI for Turkish, Korean, Chinese and Indian Armies

III. a moral case against (Marten's clause)

the decision to kill should not be delegated to a machine

"being killed by a machine is the ultimate human indignity"

Maj. Gen. Latiff

IV. 10 risks to global security

1. profliferation
2. lowered threshold for conflict
3. continuous global battlefield
4. accelerating the pace of battle
5. unpredictable interaction
6. accidental conflict
7. cyber vulnerability
8. militarisation of the civilian world
9. automated oppression
10. non-state actors

defensive systems - supervised autonomy (?)

A way forward

new york meeting october 2012

prohibition

CCW

Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May be Deemed to be Excessively Injurious or to Have Indiscriminate Effects (eg blinding laser weapons, chemical and biological weapons)

The convention has five protocols:

- **Protocol I restricts weapons with non-detectable fragments**
- **Protocol II restricts landmines, booby traps**
- **Protocol III restricts incendiary weapons**
- **Protocol IV restricts blinding laser weapons (adopted on October 13, 1995)**
- **Protocol V sets out obligations and best practice for the clearance of explosive remnants of war, adopted on November 28, 2003 in Geneva**

Conclusions 1

Autonomous Weapons Systems (AWS)

IHL compliance with AWS cannot be guaranteed for the foreseeable future.

The predictability of AWS to perform mission requirements cannot be guaranteed.

The unpredictability of AWS in unanticipated circumstances makes weapons reviews extremely difficult or even impossible to guarantee IHL compliance.

The threats to global security are unacceptably high

Conclusions 2

We are at a choice point in history where the decisions we make about automating warfare will determine the future of security.

Mass proliferation could see the full automation and dehumanisation of warfare

Let us maintain meaningful human control over the application of violent force

What can the machine learning community do?

icrac.net
responsiblerobotics.org

thank you for listening

@StopTheRobotWar
@noelsharkey