

Defining the human role in the use of force: IHL aspects

Berlin Webcast Forum on LAWS
1-2 April 2020

Kathleen Lawand
Head, Arms Unit, Legal Division, ICRC



ICRC

Key points

1. Limits on autonomy in weapon systems already exist under IHL rules governing the conduct of hostilities
 - What are those limits?
 - Does IHL already prohibit certain autonomous weapons?
2. Many questions left unanswered by existing IHL
3. What types of constraints on autonomous weapons may be required by IHL to retain sufficient / meaningful level of human control?
 - Constraints on targets and tasks
 - Spatial and temporal limits
 - Ability to supervise and intervene



Existing limits under IHL

- **IHL rules on the conduct of hostilities** – notably **distinction, proportionality and precautions in attack** – **must be complied with by those persons who plan, decide on and carry out attack.**
- The assessments required by IHL rules involve **evaluative and contextual judgements**, for which **humans are responsible and accountable.**
 - These context-based human judgments cannot be substituted with machine, sensor or software functions.
- Under existing IHL, those who use autonomous weapons must **retain a level of human control that allows them to make the required context-specific judgments.**



Existing limits under IHL

IHL rules of distinction, proportionality, precautions in attack

- e.g. direct attacks only at “military objectives”
 - “military objectives are limited to those objects which by their nature, location, purpose or use make an effective contribution to military action and whose partial or total destruction, capture or neutralization in the circumstances ruling at the time, offers a definite military advantage.” (Art. 52.2 AP I)
- e.g. Proportionality:
 - Prohibition to attack a military objective when the expected incidental civilian casualties and damage to civilian objects, or a combination thereof, would be excessive in relation to the concrete and direct military advantage anticipated. (Art. 51.5.b AP I)



Existing limits under IHL

- The legal assessments required by IHL rules demand **context-based value judgements** by those who plan, decide and carry out attack, i.e. by those humans responsible for complying with IHL in the conduct of hostilities.
- IHL requires commanders to have **knowledge** of the context, i.e. of circumstances prevailing at the time of the attack – the specific target, its location, and the time of the attack – and to be able to **predict** the weapon's effects.
- **An autonomous weapon that is unsupervised, unpredictable and unconstrained in time and space would be unlawful under IHL.**



Key questions with no clear answers in IHL

- What level of human control allows the commander / operator to exercise the context-specific judgments required by IHL?
 - What is the minimum level of predictability and reliability of the weapon system in its environment of use?
 - What constraints are needed for tasks, targets, operational environments, time of operation, and geographical scope of operation?
 - What level of human supervision, intervention and ability to deactivate is needed to comply with IHL rules?



Three types of measures and constraints

- **Constrain the targets and tasks of the autonomous weapon**
 - Not a panacea, especially in complex environments (e.g. concentrations of civilians) : combine with other constraints !
- **Constraints on the environment and situation of use, including spatial and temporal limits**
 - To ensure that planning assumptions (legal assessments) remain valid throughout the attack.
- **Provide ability for human commander / user to supervise and intervene in the operation of the autonomous weapon during the attack.**
 - Dynamic nature of most contexts would require constant supervision and communication after activation
- **In practice today, existing weapons with autonomy in their critical functions operate under all of these constraints!**



Conclusions

- Many open questions about IHL limits on autonomous weapons
- This underscores need for internationally agreed limits on autonomy in weapon systems
- Any agreed limits must build upon and strengthen existing IHL
- Focus on Guiding Principles **c** (human-machine interaction) and **d** (accountability)
- Urgency to act !

